PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 28 May 15, 1942 Number 5
Copyright 1942 by the National Academy of Sciences

RADIOACTIVE CARBON AS AN INDICATOR OF CARBON DIOXIDE UTILIZATION. IX. THE ASSIMILATION OF CARBON DIOXIDE BY PROTOZOA

By C. B. van Niel, J. O. Thomas, S. Ruben and M. D. Kamen

The Hopkins Marine Station (of Stanford University), Pacific Grove, California, the Chemical Laboratory, and the Radiation Laboratory of the University of California

Communicated March 23, 1942

Previous publications from these laboratories have demonstrated the widespread occurrence of carbon dioxide utilization by living cells.[1-7] A survey of the known cases[8] has made it extremely likely that in all cases the absorption leads primarily to the formation of a carboxylic acid in which the carbon dioxide molecule becomes incorporated as a carboxyl group.

So far, the class of protozoa had not yet been investigated in this respect. It is, however, noteworthy that Rahn[9] has recently published evidence for the absolute necessity of CO_2 in the development of these organisms, a fact which is entirely in line with the general conclusions regarding the rôle of this compound in metabolism, expressed in the previous paper of this series.[8]

Experiments by one of us (J. O. T.) with pure cultures of the holotrichous ciliate, *Tetrahymena geleii*, have shown that under anaerobic conditions a sugar fermentation occurs in which the main metabolic products are lactic, acetic and succinic acids, accompanied by carbon dioxide. Furthermore, it was ascertained that suspensions of the protozoan in a phosphate-bicarbonate buffer can take up measurable amounts of carbon dioxide. In view of the work of Wood and Werkman,[10-13] Carson and Ruben[2, 3] and others, it seemed probable that this assimilation would be responsible for the production of succinic acid. Since it is difficult, if not impossible, due to the net evolution of CO_2 to demonstrate conclusively the participation of carbon dioxide in the succinic acid formation we have studied the anaerobic metabolism of *Tetrahymena geleii* in the presence of radioactive $C^{11}O_2$.*

The ciliate was grown in yeast extract media with 2% glucose under conditions of adequate aeration. The cells were centrifuged for use in the experiments, washed and resuspended in phosphate buffer (pH 7.5) with 1%

glucose. The suspension was then shaken in the presence of $C^{11}O_2$ in an oxygen-free atmosphere for 30 minutes at 30°C. Following the addition of small amounts of lactic, acetic, pyruvic, succinic and fumaric acids as carriers the suspension was rapidly boiled in order to stop further metabolism, and to drive off dissolved CO_2. Complete elimination of residual $C^{11}O_2$ was, as usual, accomplished by addition of $NaHCO_3$, acidification and further boiling. Hereafter the suspensions were centrifuged, the cells resuspended in H_2O and again thrown down. Measurements of the insoluble cellular material at this time showed it to be practically without radioactivity.

Pyruvic and fumaric acids were precipitated from separate aliquots of the original supernatant solution with 2,4-dinitrophenylhydrazine and mercurous nitrate, respectively. These precipitates, also, showed but little activity.

From another aliquot of the original supernatant the volatile acids were removed by a vacuum distillation, and the activities of the distillate and the residue measured separately. The latter fraction contains the known products, succinic and fumaric acids.

The relative activities of the various fractions are assembled in table 1.

TABLE 1

RELATIVE ACTIVITIES OF THE FRACTIONS OBTAINED FROM A SUSPENSION OF *Tetrahymena geleii*, INCUBATED FOR 30 MINUTES IN THE PRESENCE OF $C^{11}O_2$

FRACTION	ACTIVITY IN % OF TOTAL C^*O_2 REDUCED[a]
Fumaric acid	< 4
Pyruvic acid	< 0.2
Volatile acid	< 0.1
Non-volatile residue	98
Cell material	2

[a] All values corrected for decay, hence strictly comparable.

The uptake of $C^{11}O_2$ was surprisingly large; during the incubation period about 35% of the available supply had been reduced by the protozoa. It even reached 50% in a subsequent experiment. It is important to note that the two principal products of the fermentation, lactic and acetic acid, were completely inactive, and only a small fraction of the activity was contained in the cell material after one washing with H_2O. The non-volatile residue, containing the succinic and fumaric acids, accounted for practically all the activity. Boiling with acid permanganate reduced this activity only slightly. It may be pointed out that the activity of the fumaric acid fraction may have been caused partly by co-precipitated mercurous succinate.

These results make it probable, therefore, that the $C^{11}O_2$ had actually been used in the production of succinic acid. For a convincing demon-

stration a characterization of the nature of the radioactive substance was carried out in an additional experiment in the following manner.

The protozoa suspension, prepared in the same manner as before, and after having been shaken in the presence of $C^{11}O_2$ for 30 minutes at 30°C., was again acidified, boiled and centrifuged. The supernatant, upon the addition of about 20 mg. carrier succinic acid, was neutralized (final pH 7.17), and subjected to the action of a succino-dehydrogenase preparation from beef heart in the presence of air.† The dehydrogenation of the succinic acid was followed manometrically, and was complete in 30 minutes at 37°C. The enzyme preparation was then precipitated by the addition of trichloroacetic acid, and the precipitate removed by centrifugation. The supernatant, which now contained the original succinic acid in the form of fumaric acid, was divided into three aliquots which were used for the following tests:

1. The total radioactivity was measured in one aliquot directly. It showed that 50% of the available $C^{11}O_2$ had been assimilated.

2. In the second aliquot the fumaric acid was precipitated with mercurous nitrate, the precipitate filtered, washed, dried and its activity determined.

3. The third fraction was treated with $KMnO_4$ in 1.5 N H_2SO_4 at 40°C. in a stream of nitrogen. The gas evolved was passed through $Ba(OH)_2$ solution, in which a copious precipitate of $BaCO_3$ occurred. This also was filtered, washed, dried and weighed, and its activity was measured on weighed portions. With the solution remaining after the permanganate treatment a steam distillation was performed, and the distillate, after having been made alkaline, measured in the Geiger counter.

The oxidation of fumaric acid under the conditions of the test proceeds in accord with the equation

$$\begin{array}{l} \mathrm{COOH} \\ | \\ \mathrm{CH} \\ \| \\ \mathrm{CH} \\ | \\ \mathrm{COOH} \end{array} + 2\mathrm{MnO_4^-} + 6\mathrm{H^+} \longrightarrow 3\mathrm{CO_2} + \mathrm{HCOOH} + 2\mathrm{Mn^{++}} + 4\mathrm{H_2O}$$

It has been demonstrated (Allen and Ruben[15]) that the formic acid arises exclusively from the methine carbons of the fumaric acid. Thus, a comparison of the activities of the oxidation products, CO_2 and formic acid, permits of an unambiguous localization of the C^{11} atoms in the fumaric, and, therefore, in the original succinic acid. Table 2 presents the results.

From the first experiment (table 1) it has emerged that the amount of radioactive fumaric acid in the fermentation solution is certainly not above 4%. The very large quantity of this substance in a solution subjected to

the highly specific action of succino-dehydrogenase thus shows conclusively that at least 83% of the $C^{11}O_2$ reduced had been converted by *Tetrahymena geleii* into succinic acid. This value must be considered as a lower limit because unfortunately the quantity of mercurous fumarate was large, so that considerable self-absorption of the C^{11} positrons seriously interfered with an accurate radioactivity measurement. This is borne out by the fact that the determination of C^{11} in the $BaCO_3$, carried out with a smaller amount (thinner sample), yielded a higher relative activity, undoubtedly the result of a decreased self-absorption.

A comparison of the values for the carbonate and formic acid fractions clearly proves that the labeled carbon was present in the fumaric acid only in the carboxyl groups. Together with the evidence given above, this implies that the carbon dioxide assimilated during the anaerobic sugar metabolism of *Tetrahymena geleii* is converted mainly into succinic acid, and is present exclusively in the carboxyl groups of this substance.

TABLE 2

RELATIVE ACTIVITIES OF THE VARIOUS FRACTIONS AND ALIQUOTS

FRACTION	ACTIVITY IN % OF TOTAL C^*O_2 REDUCED
Fumaric acid as mercurous fumarate	83
CO^2 produced during $KMnO^4$ oxidation of second aliquot	89
Formic acid fraction of fumarate oxidation	1

Since the mechanism for the formation of succinic acid with the above properties has been adequately dealt with elsewhere[8, 12, 13] it need not be further discussed at this place.

We are indebted to Professor E. O. Lawrence and members of the Radiation Laboratory for making these experiments possible.

Summary.—Experiments with radioactive carbon dioxide have shown that the formation of succinic acid during the anaerobic sugar metabolism of the holotrichous ciliate, *Tetrahymena geleii*, involves the assimilation of carbon dioxide. Since the radioactivity of the succinic acid so produced is strictly limited to the carboxyl groups, the results support the current concepts of the mechanism of succinic acid formation.

* The half-life of C^{11} is 20.5 minutes; the production and quantitative determination of C^{11} have been described elsewhere.[14]

† We are indebted to Mr. Bartley Carden for the enzyme preparation and for directions concerning its use.

[1] Ruben, S., and Kamen, M. D., *Proc. Nat. Acad. Sci.*, **26**, 418 (1940).

[2] Carson, S. F., Ruben, S., Kamen, M. D., and Foster, J. W., *Ibid.*, **27**, 475 (1941).

[3] Carson, S. F., and Ruben, S., *Ibid.*, **26**, 422 (1940).

[4] Barker, H. A., Ruben, S., and Kamen, M. D., *Ibid.*, **26**, 426 (1940).

[5] Barker, H. A., Ruben, S., and Beck, J. V., *Ibid.*, **26**, 477 (1940).

[6] Carson, S. F., Foster, J. W., Ruben, S., and Barker, H. A., *Ibid.*, **27**, 229 (1941).

[7] Foster, J. W., Carson, S. F., Ruben, S., and Kamen, M. D., *Proc. Nat. Acad. Sci.*, **27**, 590 (1941).

[8] van Niel, C. B., Ruben, S., Carson, S. F., Kamen, M. D., and Foster, J. W., *Ibid.*, **28**, 8 (1942).

[9] Rahn, O., *Growth*, **5**, 197 (1941).

[10] Wood, H. G., and Werkman, C. H., *Biochem. Jour.*, **30**, 48 (1936).

[11] Wood, H. G., and Werkman, C. H., *Ibid.*, **32**, 1262 (1938).

[12] Wood, H. G., and Werkman, C. H., *Ibid.*, **34**, 7, 129 (1940).

[13] Wood, H. G., and Werkman, C. H., Hemingway, A., and Nier, A. O., *Jour. Biol. Chem.*, **139**, 483 (1941).

[14] Ruben, S., Kamen, M. D., and Hassid, W. Z., *Jour. Am. Chem. Soc.*, **62**, 3443 (1940).

[15] Allen, M. B., and Ruben, S., *Ibid.* (in press).

THE ENERGY EQUATION FOR A VISCOUS COMPRESSIBLE FLUID

BY H. J. STEWART

GUGGENHEIM AERONAUTICS LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated April 6, 1942

One of the more important problems discussed in any text book on Dynamic Meteorology is the calculation of the energy of atmospheric motions, and the conversion of potential and thermal energy to the kinetic energy of the motion of the winds. This problem was first discussed by Margules[1] and has since then been discussed by Bjerknes,[2] Brunt[3] and Haurwitz.[4] Rossby has also discussed this work in his lectures on Dynamic Meteorology. It is therefore of considerable interest to note that the development of the energy equation for a viscous compressible fluid by all these latter writers is in error through the omission of terms involving the viscous stresses. On the other hand, this matter is correctly handled in Margules' original paper.

The energy equation for this case can be obtained easily from the dynamical equations of motion for a viscous fluid. This is, using the Cartesian tensor notation,

$$\frac{Du_i}{Dt} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} - \frac{\partial V}{\partial x_i} + \frac{1}{\rho}\frac{\partial}{\partial x_j}(\tau_{ij}) \qquad (1)$$

where x_i are rectangular coördinates, u_i is the fluid velocity vector, ρ is the fluid density, V is the gravitational potential and τ_{ij} is the viscous stress tensor. Upon taking the scalar product of the velocity vector with this, one obtains, since $\dfrac{\partial V}{\partial t} = 0$,

$$\rho \frac{D}{Dt}\left\{\frac{1}{2}u_i u_i + V\right\} = -u_i \frac{\partial p}{\partial x_i} + u_i \frac{\partial}{\partial x_j}(\tau_{ij}). \tag{2}$$

If heat is added to the fluid, the relative energy equation for the element is

$$\frac{Dq}{Dt} = \frac{Di}{Dt} + p\frac{D}{Dt}\left(\frac{1}{\rho}\right) - \frac{1}{\rho}\tau_{ij}\frac{\partial u_i}{\partial x_j} \tag{3}$$

where q is the heat added per unit mass of fluid, and i is the internal energy of a unit mass of fluid. The last two terms represent the work done on the external fluid through the deformation of the element. The equation of continuity,

$$0 = \frac{1}{\rho}\frac{D\rho}{Dt} + \frac{\partial u_i}{\partial x_i}, \tag{4}$$

can be used to transform equation (3) to

$$\rho\frac{Dq}{Dt} = \rho\frac{Di}{Dt} + p\frac{\partial u_i}{\partial x_i} - \tau_{ij}\frac{\partial u_i}{\partial x_j}. \tag{5}$$

In order to obtain the complete energy equation, equations (3) and (5) must be added. This gives the energy equation for a viscous compressible fluid as

$$\rho\frac{Dq}{Dt} = \rho\frac{D}{Dt}\left\{i + V + \frac{1}{2}u_i u_i\right\} + \frac{\partial}{\partial x_i}(pu_i) - \frac{\partial}{\partial x_j}(\tau_{ij}u_i). \tag{6}$$

This equation could have been written directly by the principle of conservation of energy as the last two terms are the work done by the element on the surrounding fluid by pressure and viscous stresses, respectively.

In each of the above mentioned references, the term, $\tau_{ij}\frac{\partial u_i}{\partial x_j}$, is omitted from equation (3), and the final energy equation is written as

$$\rho\frac{Dq}{Dt} = \rho\frac{D}{Dt}\left\{i + V + \frac{1}{2}u_i u_i\right\} + \frac{\partial}{\partial x_i}(pu_i) - u_i\frac{\partial}{\partial x_j}(\tau_{ij}). \tag{7}$$

That this equation is incorrect is easily seen by considering its application to a closed system with fixed walls and to which no heat is added. An auxiliary theorem will first be developed for use in this demonstration. By use of the continuity equation, it can easily be seen that

$$\rho\frac{DF}{Dt} = \frac{\partial}{\partial t}(\rho F) + \frac{\partial}{\partial x_i}(\rho F u_i) \tag{8}$$

where F is any given quantity.

Thus,

$$\int\int\int_V \rho \frac{DF}{Dt}\, dv = \frac{\partial}{\partial t} \int\int\int_V \rho F dv \tag{9}$$

with the volume integral extending throughout the given closed system. The integral of the second term of equation (8) vanishes since $u_i = 0$ on the boundaries. By applying this result to equation (7), it is seen that

$$\frac{\partial}{\partial t}\{I + P + K\} = -\int\int\int_V \frac{\partial}{\partial x_i}(pu_i)dv + \int\int\int_V u_i \frac{\partial}{\partial x_j}(\tau_{ij})dv \tag{10}$$

where I, P and K are the total internal, potential and kinetic energies within the system. The first integral on the right-hand side vanishes, and the second can be transformed by integration by parts to

$$\frac{\partial}{\partial t}\{I + P + K\} = -\int\int\int_V \tau_{ij} \frac{\partial u_i}{\partial x_j}\, dv. \tag{11}$$

Now, $\tau_{ij}\dfrac{\partial u_i}{\partial x_j}$ is the kinetic energy dissipated per unit time and volume and turned into heat, so this would indicate that the total energy within the system is diminishing at a rate equal to the rate of dissipation of kinetic energy. This result is in contradiction to the principle of the conservation of energy, and equation (7) must therefore be incorrect. Equation (10) is given by Haurwitz[4] (p. 245) although its significance is not recognized.

If the corect energy equation, equation (6), is applied to the given closed system, it is seen that

$$\frac{\partial}{\partial t}\{I + P + K\} = 0. \tag{12}$$

By the principle of the conservation of energy this is obviously the correct result.

Since $\tau_{ij}\dfrac{\partial u_i}{\partial x_j}$ is the kinetic energy converted into heat per unit time and volume, equation (3) can be changed into another form,

$$\frac{Dq'}{Dt} = \frac{Dq}{Dt} + \frac{1}{\rho}\tau_{ij}\frac{\partial u_i}{\partial x_j} = \frac{Di}{Dt} + p\frac{D}{Dt}\left(\frac{1}{\rho}\right), \tag{13}$$

where $\dfrac{Dq'}{Dt}$ is the total heat added to the element with one part, $\dfrac{Dq}{Dt}$, being

added from outside and the other part, $\frac{1}{\rho}\tau_{ij}\frac{\partial u_i}{\partial x_j}$, being an internal heat source arising from the viscous dissipation. This result has been noted by Lamb[5] (p. 646). The error in equation (7) can thus be considered as arising from neglecting the heat produced by viscous dissipation of the kinetic energy.

[1] Margules, M., *Met. Z.*, **23**, 481 (1906).

[2] Bjerknes, V., and collaborators, *Physikalische Hydrodynamik*, Chap. 4, Julius Springer, Berlin, 1933.

[3] Brunt, D., *Physical and Dynamical Meteorology*, Chap. XV, The MacMillan Company, New York, 1934.

[4] Haurwitz, B., *Dynamic Meteorology*, Chap. XII, McGraw-Hill Book Co., Inc., New York, 1941.

[5] Lamb, H., *Hydrodynamics*, 6th ed., Cambridge University Press, Cambridge, 1932.

SENSITIZATION OF MELANOPHORES BY NERVE CUTTING

By G. H. Parker

Biological Laboratories, Harvard University

Communicated April 7, 1942

For some years past evidence has been accumulating to show that effectors and even neurones on denervation became more sensitive to their ordinary chemical activators than they were before their special nervous connections were severed. This question has been discussed recently by Cannon and Rosenblueth (1937) and has been made the subject of a special lecture by Cannon (1939) who has suggested that such sensitization might well represent a law of denervation. The evidence upon which this view is based has been gathered from the responses of denervated smooth muscles, glands, skeletal muscles and isolated neurones to their normal activators. Within a year similar evidence from another category of effectors, the chromatophores, has been presented by Smith (1941) who has shown that the melanophores of a fish, the tautog, can also be sensitized by denervation.

Recent studies on the color changes in catfishes have led me to suspect that the melanophores of this creature may also be open to sensitization by nerve cutting and during the last year I have carried out a set of tests to ascertain if this were true or not. The work was begun at the Marine Biological Laboratory, Woods Hole, and completed at the Harvard Biological Laboratories. To both these institutions I am indebted for accommodations and ample facilities. The cost of materials, reagents and the like used in these investigations was met in large part by a grant from the Permanent Science Fund to the officers of which I wish to express my sincere thanks.

Smith's procedure with the tautog was to determine the times necessary for the concentration of pigment in three states of the color cells of this fish when activated by adrenaline. Melanophores were tested in freshly removed scales with the normal nerve-terminals still attached, in scales with fully degenerated terminals and in scales with regenerated terminals. The result of Smith's investigations was to show that the pigment in cells with degenerated terminals became concentrated about twice as fast as in cells with active terminals.

My own method in testing this question was very different from that used by Smith, and in place of a marine fish with scales like the tautog I used a scaleless, fresh-water fish, the common eastern catfish, *Ameiurus nebulosus*. The chromatophore system in this fish consists exclusively of melanophores, micromelanophores in the epidermis and macromelanophores in the derma. Of these the macromelanophores are much more satisfactory for study and are the color cells that will be referred to in the present investigation. The melanophore pigment in *Ameiurus* is concentrated by the action of nerve-fibres whose effective neurohumor appears to be adrenaline. It is dispersed by two agents, intermedine from the pituitary gland and acetylcholine from an appropriate set of nerve-fibres (Parker, 1940). These facts will be of importance in the further steps of this investigation.

The responses of denervated and innervated melanophores in *Ameiurus* to adrenaline can be conveniently studied in caudal bands on this fish. When a single fin ray is cut in the tail of *Ameiurus* to form a caudal band nerves carrying both dispersing and concentrating fibres are severed. After having been cut the severed fibres remain active for several days, in two weeks they are completely degenerated and in about a month they have regenerated. (Parker, 1934, 1941*a*; Abramowitz, 1935, 1936.) The best period in which to obtain a band of denervated melanophores is in the third week after the severance of the nerves. All these time relations apply to fishes during the summer when the water in which they live is at approximately 20°C. The adrenaline used in these tests was the Parke, Davis & Co. preparation 1:1000 which was diluted with cold-blooded Ringer's solution to the requisite strength (see table 1). The catfishes tested weighed each about 50 g. and received by subcutaneous injection 0.2 cc. of the given concentrations of adrenaline. At the beginning of the tests the fishes were intermediate in tone. The effects on the innervated tails and on their denervated bands are given in the third and fourth columns of table 1. These effects are best seen in the first five minutes after the injection has been made. Strong concentrations of adrenaline (1:5000 and 1:20,000) blanched the fishes quickly, the tails and bands becoming pale simultaneously. The injections of weak concentrations (1:1,000,000 to 1:50,000,000) had no appreciable effect upon the tails and bands within the first

five minutes. At a concentration of 1:1,000,000 the whole fish including the band and tail commonly blanched markedly in about an hour. At greater dilutions, however, no change whatever in tint was seen.

TABLE 1

COLOR RESPONSES OF THE INNERVATED TAILS (AS WELL AS THE BODIES OF THE FISHES) AND OF THE DENERVATED CAUDAL BANDS OF CATFISHES (*Ameiurus*) DURING THE FIRST FIVE MINUTES AFTER THE INJECTION OF DIFFERENT CONCENTRATIONS OF ADRENALINE

Three fishes of intermediate tint were used in each test and each fish, which weighed about 50 g., received 0.2 cc. of the solution of adrenaline. The amounts of adrenaline in milligrams injected per 100 g. of fish are given in the second column of the table. The concentrations were made by diluting with Ringer's solution the Parke, Davis & Co. preparation of adrenaline chloride 1:1000. (1) Palish after 10 minutes. (2) Palish after 15 minutes. (3) Became intermediate in an hour. (4) Became pale in an hour. (5) No change whatever.

CONCENTRATIONS OF ADRENALINE	ADRENALINE MG. PER 100 G. FISH	TINT OF REGIONS IN FIRST 5 MINUTES	
		INNERVATED TAIL	DENERVATED BAND
1:5,000	0.08	Blanched quickly	Blanched quickly
1:20,000	0.02	Blanched quickly	Blanched quickly
1:50,000	0.008	Intermediate (1)	Decidedly pale
1:100,000	0.004	Intermediate (2)	Somewhat pale (3)
1:1,000,000	0.000,4	Intermediate (4)	Intermediate (4)
1:20,000,000	0.000,02	Intermediate (5)	Intermediate (5)
1:50,000,000	0.000,008	Intermediate (5)	Intermediate (5)

The critical concentrations in this set of records were those of medium strength (1:50,000 and 1:100,000). In the stronger of these the tail remained intermediate in tint while the band became decidedly blanched. The melanophore pigment masses in the tail generally were stellate, but in the band they were irregularly roundish. About ten minutes after the injection had been made the pigment masses of the tail melanophores became generally roundish. With the weaker concentration 1:100,000 the tail during the first five minutes remained intermediate but the bands blanched noticeably. In about an hour the bands were again intermediate. Thus of the seven concentrations of adrenaline recorded in table 1 two (1:50,000 and 1:100,000) were effective in the first five minutes in concentrating the pigment in the denervated melanophores but failed to do so in the innervated ones.

Two other sets of catfishes were tested for adrenaline responses on essentially the same plan as that outlined in table 1. In both these sets the concentrations began at 1:5000; in one they ran to 1:40,000,000 and in the other to 1:100,000,000. In both sets of fishes the extreme concentrations called forth similar responses in the tails and their associated bands. No differential responses were noted at concentrations stronger than 1:50,000, nor weaker than 1:200,000. Five of the six fishes tested with concentration 1:50,000 responded differentially to this solution as did those recorded in table 1. In the sixth fish both the band and tail blanched indistinguish-

ably. To the concentration 1:200,000 (0.002 mg. of adrenaline per 100 g. of fish) the bands in three fishes responded slightly by advanced blanching thus increasing the range of this effect beyond what is shown in table 1. Aside from these minor differences, however, all three sets of tests were in reasonable agreement in that extreme concentrations of adrenaline either strong or weak influenced the melanophores both denervated and innervated in the same way, whereas the intermediate concentrations (1:50,000 to 1:200,000) called forth a differential response in which the denervated color cells were seen to be more sensitive to adrenaline than the innervated ones. Thus these tests on *Ameiurus* support Smith's results (1941) on the sensitization of the melanophores of the tautog by denervation and afford among chromatophores another example of this type of response.

In addition to adrenaline two other important neurohumors, as already mentioned, take part in the activation of catfish melanophores, acetylcholine and intermedine. Both these are darkening agents. Sensitization of smooth muscle to acetylcholine by nerve cutting has already been noted (Cannon and Rosenblueth, 1937) and it is therefore possible that this phenomenon might be met with in *Ameiurus* melanophores. It was this possibility in fact that led me to suggest (Parker, 1941*b*) what might be an explanation of the sensitization of denervated catfish melanophores to adrenaline. When a caudal band is cut in a catfish both cholinergic and adrenergic fibres are severed and finally degenerate. Consequently in time such a band would become deficient in the two opposing neurohumors, acetylcholine and adrenaline. In the absence of acetylcholine, therefore, the denervated melanophores would be more freely open to the action of adrenaline than when its opponent neurohumor was present. Hence denervation by reducing acetylcholine might render the melanophores more sensitive to adrenaline. Acetylcholine, however, is limited in its action on catfish melanophores (Osborn, 1938; Parker, 1940). It can convert the pale phase into the intermediate one but not the intermediate one into the dark. Thus it acts over only about half the range of the melanophore change.

Notwithstanding this limitation acetylcholine was tested on palish catfishes with bands that had been previously eserinized (Parker, 1940). The concentrations of acetylcholine used ran from 1:1000 to 1:100,000,000 and each fish received 0.2 cc. of the solution with which it was tested. The weaker concentrations were without effect and the stronger ones proved poisonous as was evidenced by the death of many of the fishes. Only at 1:100,000 or 0.004 mg. of acetylcholine per 100 g. of fish did the fishes darken and continue to live. At this concentration, which is about one-hundredth of that used by Kiel and Root (1941) in tests on smooth muscle in the cat, no noticeable differential response could be detected between the bands and the rest of the tail. This occurred so regularly that I was led to

conclude that either acetylcholine was a substance to which the melanophores could not be sensitized by denervation or that the technique herein employed was insufficient to bring out this capacity if it was present.

The second darkening neurohumor in the catfish is the pituitary product, intermedine. The injection of aqueous extracts of the pituitary gland of the catfish are well known to darken this animal (Parker, 1934; Abramowitz, 1936). This fish will also darken to injections of Parke, Davis & Co.'s preparation of obstetrical pituitrin which contains intermedine. Groups of three catfishes of intermediate tint and with fully denervated caudal bands were injected each with a given concentration of pituitary extract made from the glands of the catfish itself. These glandular extracts were made by grinding to great fineness in a known volume of Ringer's solution a definite number of pituitary glands taken freshly from the fishes, and then diluting the mixture till the requisite degree was reached. The volume of extract injected in each fish was 0.2 cc. and the concentration of pituitary tissue in each dose injected in terms of glands varied from two to one-twentieth. After the injections were made the bands and tails of the fishes were watched closely for the next few hours. The significant changes began to show within a very short time. When the three fishes, intermediate in tint, were injected with the most concentrated extract, two glands per dose, all began to darken within five minutes. At ten minutes their bands were noticeably darker than the rest of their tails. This condition remained constant for from three to four hours after which both bands and tails were indistinguishably dark and of the same deep tint as the rest of the fish. This condition was maintained for fully a day whereupon the fishes gradually blanched from having been kept in a white, illuminated vessel. When the concentration injected was one gland per fish the changes were essentially the same as those observed with the strongest extract. For the first few hours after injection the bands were clearly darker than their surroundings after which all were of the same deep tone. When the concentration was half a gland per dose the fishes darkened slowly. For about an hour the bands darkened more rapidly than the tails. Later both bands and tails were equally dark. When only a tenth of a gland was injected no certain difference could be distinguished between the bands and the tails, both of which, however, darkened. After injecting one-twentieth of a gland into each fish no change of any kind was noticeable but all three fishes remained of intermediate tint in close agreement with a group of uninjected, intermediate catfishes held as a check. From these records it appears clear that the denervation of catfish melanophores renders them more sensitive to their own intermedine than they were when innervated and that this newly acquired state is comparable to that of adrenaline sensitization after nerve cutting. Experiments of much the kind as those detailed in this paragraph have already been carried out

by other workers (von Frisch, 1911; Abramowitz, 1936), but these investigators were more interested in the final darkening of the fishes than in the steps by which this concluding stage was attained and consequently gave no attention to the changes that have proved significant in the present research.

Other ways of subjecting catfishes with innervated and denervated melanophores to their own intermedine than the one here described were also tried. It is well known that enucleation will induce catfish in the light to become very dark through a discharge of intermedine. Fishes with bands of appropriate age were therefore deprived of their eyes and kept in the light. They quickly darkened in color but the bands darkened indistinguishably from the tails. The failure of the bands to show an advanced darkening was probably due to the rapidity with which under these circumstances the dark phase was attained. I also failed to obtain darker bands than tails by using obstetrical pituitrin as a dispersing agent. After the injection of this extract the fishes darkened but not in a differential way. This extract which is mammalian in source is very probably less reliable as a dispersing agent than the fresh one made directly from the pituitary glands of the catfishes themselves. When hypophysectomized catfishes with seasoned caudal bands were injected with a pituitary extract of one-gland strength they darkened as normal fishes with pituitary glands did, in that their caudal bands darkened more deeply than the rest of the tail and in advance of it. This is what should be expected if the account of the color reactions of the catfish thus far given is correct. From the results here detailed it appears that denervation sensitizes catfish melanophores to intermedine in much the same way that it sensitizes these color cells to adrenaline.

Although catfish melanophores can be sensitized by nerve cutting to intermedine, the fact that they cannot be so sensitized to acetylcholine is strong evidence against the explanation of melanophore sensitization that I have offered elsewhere (Parker, 1941*b*). That explanation depended on the local opposition of one neurohumor to another, acetylcholine to adrenaline. But now that it has been shown that acetylcholine is apparently not involved in sensitization this explanation, even limited as it is to fish melanophores, lacks application. Melanophore sensitization due to nerve cutting like that seen in other sensitizable effectors is so far as I know without obvious explanation.

Summary.—1. Denervation of melanophores in the catfish *Ameiurus nebulosus* will sensitize these color cells to appropriate strengths of adrenaline (0.002 to 0.008 mg. of adrenaline per 100 g. of fish). This result supports Smith's conclusion concerning the sensitization of the melanophores in the tautog.

2. Denervated catfish melanophores were not found to be sensitized to acetylcholine.

3. Catfish melanophores were sensitized by nerve cutting to intermedine whose strength of injection was two pituitary glands to half a gland per fish. These were not sensitized to weaker extracts.

4. No general explanation was found for effector sensitization by nerve cutting.

Abramowitz, A. A., "Regeneration of Chromatophoral Nerves," *Proc. Nat. Acad. Sci.*, **21**, 137–141 (1935).

Abramowitz, A. A., "Physiology of the Melanophore System in the Catfish, *Ameiurus*," *Biol. Bull.*, **71**, 259–281 (1936).

Cannon, W. B., "A Law of Denervation," *Amer. Jour. Med. Sci.*, **198**, 737–750 (1939).

Cannon, W. B., and Rosenblueth, A., *Autonomic Neuro-Effector Systems*, New York, 1937, 229 pp.

von Frisch, K., "Beiträge zur Physiologie der Pigmentzellen in der Fischhaut," *Arch. ges. Physiol.*, **138**, 319–387 (1911).

Keil, F. C., and Root, W. S., "Parasympathetic Sensitization in the Cat's Eye," *Amer. Jour. Physiol.*, **132**, 437–445 (1941).

Osborn, C. M., "The Role of the Melanophore-Dispersing Principle of the Pituitary in the Color Changes of the Catfish," *Jour. Exp. Zool.*, **79**, 309–330 (1938).

Parker, G. H., "Color Changes of the Catfish *Ameiurus* in Relation to Neurohumors," *Ibid.*, **69**, 199–233 (1934).

Parker, G. H., "On the Neurohumors of the Color Changes in Catfishes and on Fats and Oils as Protective Agents for Such Substances," *Proc. Amer. Philos. Soc.*, **83**, 379–409 (1940).

Parker, G. H., "Melanophore Bands and Areas Due to Nerve Cutting, in Relation to the Protracted Activity of Nerves," *Jour. Gen. Physiol.*, **24**, 483–504 (1941*a*).

Parker, G. H., "Hypersensitization of Catfish Melanophores to Adrenaline by Denervation," *Biol. Bull.*, **81**, 302 (1941*b*).

Smith, D. C., "The Effect of Denervation upon the Response to Adrenalin in the Isolated Fish Scale Melanophores," *Amer. Jour. Physiol.*, **132**, 245–248 (1941).

ON GENERALIZED BIORTHOGONAL EXPANSIONS IN METRIC AND UNITARY SPACES

BY Y. Y. TSENG

DEPARTMENT OF MATHEMATICS, TSING-HUA UNIVERSITY, KUNMING, CHINA

Communicated April 7, 1942

1. Introduction.—Recently R. P. Boas, Jr., enunciated in these PROCEEDINGS[1] a principle of biorthogonal expansion, from which follow readily various well-known special expansion theorems. The object of the present note is to give new conditions for expansions, to generalize and strengthen some of the principal results on general expansions occurring in the litera-

ture. This is done by introducing for study certain closely and directly related functional transformations and infinite matrices, with the aid of E. H. Moore's second theory of general analysis (*G. A.*)[2] and a few facts newly encountered there.

Unless specified otherwise, we consider throughout this note (1) *an arbitrarily fixed (semi-definite) positive Hermitian matrix*[2] E_* on PP, with "coördinates" $E_*(p'p'')$ where $p'p''$ run independently over the *unrestricted* non-empty set P, and (2) *a general system of functions* $\{g_p\}$ in a metric space $\mathfrak{R}$ or a unitary space $\mathfrak{M}$ (both being complete and left-linear over the real, complex or *quaternionic* numbers), where $\{g_p\}$ is *not* assumed linearly independent. It forms a biorthogonal system in our generalized sense, briefly a *B_*-system*,[3] if, relative to E_* there exists a system of bounded linear functionals $\{\psi_p\}$ such that

$$\psi_{p'}(g_{p''}) = E_*(p''p') \qquad (p'p'' \text{ in } P).$$

In the sequel it is fundamental that in any complete linear metric space the Cauchy condition on the existence of limit holds for any directed system of elements.[4]

A detailed account of these results and a discussion of milder conditions as well as of other expansions defining oblique projectors will be published elsewhere.

2. *Expansions in Metric Spaces.*—Let $\mathfrak{R}$ be a complete linear metric space, and $\{f_p\}$, $\{g_p\}$ be arbitrary systems of elements in $\mathfrak{R}$. Then $\{f_p\}$ is said to be completely accordant to $\{g_p\}$ or completely g-accordant, if, for any bounded linear functional L on $\mathfrak{R}$,

$$L(g_p) = 0\ (p) \quad \text{implies} \quad L\ (f_p) = 0\ (p)\ .$$

We call $\{f_p\}$ bounded relative to $\{g_p\}$ or g-bounded, in case there exists a positive constant M such that

$$\Big\|\sum_{p\in\sigma} a_p f_p\Big\| \leqq M \cdot \Big\|\sum_{p\in\sigma} a_p g_p\Big\|$$

for every finite subset σ of P and arbitrary constants $\{a_p\}$.

Relative to E_* we define a functional integral J_* provided the following limit exists:

$$J_* a_p f_p \equiv \lim_{\sigma} \sum_{p'\in\sigma} \sum_{p''\in\sigma} a_{p'} R_*^{\sigma}(p'p'') f_{p''},$$

where R_*^{σ} is Moore's general reciprocal[2] of the finite square matrix $E_*(p'p'')$ as $p'p''$ vary independently over σ, and the limit as σ swells is the one introduced years ago by Moore in these PROCEEDINGS.[4] Here and below we always write lim for strong limit.

By an *E_*-base* of $\mathfrak{R}$ we mean a B_*-system $\{g_p\}$ with an adjoint system $\{\psi_p\}$ in terms of which every h of $\mathfrak{R}$ has and equals the strongly con-

vergent expansion $J_*\psi_p(h)g_p$. Observe that relative to an E_*-base, the expansions are always unique with E^*-accordant[2] coefficients.

With these preliminaries we can now state:

THEOREM A. *Suppose that* $\{g_p\}$ *and* $\{f_p\}$ *are mutually bounded. Then* $\{g_p\}$ *is a completely f-accordant* E_**-base of* $\mathfrak{R}$ *if and only if* $\{f_p\}$ *is a completely g-accordant* E_**-base of* $\mathfrak{R}$.

COROLLARY (α). (Generalization of Boas's theorem.) *If for any set of arbitrary constants* $\{a_p\}$

$$\left|\left|\sum_{p\epsilon\sigma} a_p(f_p - g_p)\right|\right| \leqq \lambda \cdot \left|\left|\sum_{p\epsilon\sigma} a_p f_p\right|\right|, \quad 0 < \lambda < 1,$$

then $\{g_p\}$ *is an* E_**-base of* $\mathfrak{R}$ *whenever* $\{f_p\}$ *is such.*

COROLLARY (β). (First generalization of Paley-Wiener's theorem.[5]) *Let* $\{f_p\}$, $\{g_p\}$ *be systems in a complete unitary space, satisfying the condition of Corollary* (α). *Then* $\{g_p\}$ *is an* E_**-base, when and only when* $\{f_p\}$ *is such.*

Here, the usual proof by iteration[1, 5] breaks down, because such a positive constant λ less than one is no longer available. We note, however, that if $\{f_p\}$ is an E_*-base with an adjoint system $\{\phi_p\}$, the integral $J_*\phi_p(h)g_p$ converges strongly and defines a bounded linear transformation Ah on $\mathfrak{R}$. Moreover, it has a unique bounded classical reciprocal A^{-1} on $\mathfrak{R}$. Consequently the adjoint $(A^{-1})^*$ is also bounded. Then, defining the functionals $\{\psi_p\}$ to be $(A^{-1})^*\phi_p$, we see that $\{g_p\}$ is a B_*-system with $\{\psi_p\}$ as an adjoint system, and obtain for arbitrary h the strongly convergent expansion $J_*\psi_p(h)g_p$.

3. Expansions in Unitary Space.—When an inner product is present, sharper results and explicit formulae can be established. Hereafter, $\mathfrak{M}_g$ denotes the closure of $\{g_p\}$ (i.e., the closed linear subspace determined by these elements); P_g, the orthogonal projection on $\mathfrak{M}_g$; E_g, the Gramian matrix $(g_{p'}, g_{p''})$ of $\{g_p\}$; and $J_g a_p h_p$, the strong limit analogous to $J_* a_p g_p$ with E_*, $\{g_p\}$ replaced by E_g, $\{h_p\}$, respectively. We call the unique system $\{h_p\}$ in $\mathfrak{M}_g$ such that $(g_{p'}, h_{p''})$ equals $E_*(p'p'')$, the E_*-adjoint system to $\{g_p\}$.

With the help of new matricial results in *G. A.*, recent results on systems of functions,[6] and several lemmas some of which generalize theorems of A. J. Pell[7] and S. Lewin,[8] we derive the consequences collected in the next theorem, where, it is to be noted, the essential restriction of "nearness"[1, 5] as measured by $0 < \lambda < 1$ is entirely discarded.

THEOREM B. *In order that the* g_p*'s form a* B_**-system, the integral* $J_* c_p g_p$ *converge weakly for every* E_**-modular*[2] $\{c_p\}$, *and every g of* $\mathfrak{M}_g$ *have such a weak expansion, it is necessary and sufficient that* E_* *and* E_g *be mutually modular.*[2]

If so, the expansion is unique and converges strongly. The closures of

$\{g_p\}$ and its unique E_*-adjoint system $\{h_p\}$ coincide; E_* and E_h are mutually modular; and

$$c_p = J_g(g, g_{p'})E_*(p'p) = J_*(g, g_{p'})E_g^{-1}(p'p),$$
$$J_*(f, h_p)g_p = P_g f = J_*(f, g_p)h_p.$$

Moreover, we have the sharpened and generalized Paley-Wiener inequalities:[9,2]

$$\frac{1}{\overline{M}_*(E_g)} J_*(f, g_p)(g_p, f) \leqq ||P_g f||^2 \leqq \frac{1}{*\underline{M}(E_g)} J_*(f, g_p)(g_p, f),$$

$$\underline{M}_*(E_g) \cdot J_*(f, h_p)(h_p, f) \leqq ||P_g f||^2 \leqq \overline{M}_*(E_g) \cdot J_*(f, h_p)(h_p, f).$$

Lack of space prevents describing further expansions closely attached to the above situation. From this theorem we deduce the following result which has much contact with the recent literature.

THEOREM C. *Given two non-zero positive constants d, e with d less than e. For the existence of a system $\{f_p\}$ in $\mathfrak{M}$ such that*

(1a) $(f_{p'}, f_{p''}) = e^2 \cdot E_*(p'p'')$,

(1b) $\left|\left|\sum_{p \epsilon \sigma} a_p(g_p - f_p)\right|\right|^2 \leqq d^2 \cdot \sum_{p' \epsilon \sigma} \sum_{p'' \epsilon \sigma} a_{p'} E_*(p'p'') \overline{a_{p''}}$

it is necessary and sufficient that one of the equivalent sets of conditions (2) *and* (3) *be satisfied:*

(2a) *the rows of E_* are modular as to E_g,*

(2b) $\frac{1}{(e+d)^2} J_*(f, g_p)(g_p, f) \leqq ||P_g f||^2 \leqq \frac{1}{(e-d)^2} J_*(f, g_p)(g_p, f)$;

(3a) *E_* and E_g are mutually modular,*

(3b) $(e-d)^2 \leqq \underline{M}_*(E_g) \leqq \overline{M}_*(E_g) \leqq (e+d)^2$.

In fact, any system $\{f_p\}$ fulfilling conditions (1a) (1b) *and such that one of the spaces $\mathfrak{M}_f$, $\mathfrak{M}_g$ lies in the other must have $\mathfrak{M}_f$ coincident with $\mathfrak{M}_g$. Again, the closures of $\{g_p\}$ and its unique E_*-adjoint $\{h_p\}$ are equal. We have*

$$\underline{M}_*^2(E_g) \cdot J_g(f, h_p)(h_p, f) \leqq ||P_g f||^2 \leqq \overline{M}_*^2(E_g) \cdot J_g(f, h_p)(h_p, f).$$

Finally, the strongly convergent expansions, Paley-Wiener inequalities and other results in Theorem B all remain valid.

COROLLARY (α). (Second generalization of Paley-Wiener's theorem.) *In particular, the theorem holds when $d = \lambda < 1$, $e = 1$.*

COROLLARY (β). (Generalization of Duffin-Eachus' theorem.[10]) *Let E_* be a row-finite matrix, and $\{g_p\}$ a "minimal system" in which the omission of any term renders the closure of the remaining system a proper subspace of $\mathfrak{M}_g$. Then condition* (2b) *is both necessary and sufficient for* (1a) (1b).

Concerning this theorem and its corollaries, a few remarks may serve to clarify the situation.

Remarks.—(α) Counter-examples can be constructed to show that, unlike the ones stated in this and the preceding theorems, the inequalities as originally given by Paley and Wiener in their important closure theorem are not the best possible. (β) In the cited abstract Duffin and Eachus did not mention any additional requirement beside (2*b*). But, in any pair of systems $\{f_p\}$, $\{g_p\}$ satisfying conditions (1*a*) (1*b*) with E_* equal to the (infinite) identity matrix I, the system $\{g_p\}$ must always be minimal. And, there is a linearly independent system $\{g_p\}$ having property (2*b*), which, nevertheless, fails to be minimal even when $E_* = I$. (γ) The restriction $d < e$ is really not so narrow as it may seem. Indeed, the theorem applies to the last example of Lewin[8] and supplies more information.

As to the proof, it is in the part (2) $\rightarrow$ (1) that we find a point of special interest. Let $\{w_p\}$ be a complete system for $\mathfrak{M}_g$ with its Gramian equal to E_* and K_g the matrix whose "coördinates" $K_g(p_1p_2)$ are (g_{p_1}, w_{p_2}). Then, K_g possesses a classical right reciprocal, since $J_*K_g{}^*K_g$, $J_*K_gK_g{}^*$ each has lower E_*-modulus[2] greater than zero by virtue of a previous result on "Schmidt's problem."[6] As in the theory of linear transformations,[11] this fact in turn permits us to write K_g in its right polar form

$$K_g = J_*S_gU_g,\ S_g = \sqrt{E_g}$$

where S_g is (definite) positive Hermitian and U_g unitary. If we now set

$$f_p \equiv e \cdot J_*U_g(pp_1)w_{p_1},$$

we can show that this system has all the required properties in (1) by reference to some matricial results of *G. A.*

[1] These PROCEEDINGS, **26**, 139–143 (1940).

[2] This work of Moore, entitled *General Analysis*, is being published in *Memoirs of the American Philosophical Society*, 1935– . The less familiar notions of *G. A.* here utilized are the following: (*a*) A Hermitian matrix E_* is (semi-definite) *positive*, in case every principal minor determinant of finite order has non-negative value. (*b*) Let the canonical form of a Hermitian matrix H be $H = \sum_m r_mI_m$ where r_m is a real characteristic value of H, and I_m the idempotent Hermitian matrix associated with r_m. Then we call the following matrix the *general reciprocal* of H: $R_m \equiv \sum_m 1/r_m{}'I_m$, the prime indicating that the fraction is taken to be zero when the denominator vanishes. (*c*) We call a system of constants $\{c_p\}$ *accordant* to E_* in case

$$\sum_{p'\epsilon\sigma} \sum_{p''\epsilon\sigma} a_{p'} E_*(p'p'')\bar{a}_{p''} = 0 \quad \text{implies} \quad \sum_{p\epsilon\sigma} c_p\bar{a}_p = 0.$$

(*d*) $\{c_p\}$ is E_*-*modular* in case there exists a constant N such that

$$\left|\sum_{p\epsilon\sigma} c_p\bar{a}_p\right|^2 \leqq N \cdot \sum_{p'\epsilon\sigma} \sum_{p''\epsilon\sigma} a_{p'}E_*(p'p'')\bar{a}_{p''}$$

for every system of constants $\{a_p\}$. (e) A Hermitian matrix H is E_*-*modular*, if we can find a constant M with

$$\left|\sum_{p' \epsilon \sigma} \sum_{p'' \epsilon \sigma} \bar{a}_{p'} H(p'p'') \bar{a}_{p''}\right| \leqq M \cdot \sum_{p' \epsilon \sigma} \sum_{p'' \epsilon \sigma} a_{p'} E_*(p'p'') \bar{a}_{p''}.$$

The lower and upper *moduli*, denoted by $\underline{M}_*(H)$, $\overline{M}_*(H)$ are, respectively, the lower and upper bounds of $J_* J_* c_{p'} H(p'p'') c_{p''}$ for all E_*-modular systems $\{c_p\}$ with $J_* c_p \bar{c}_p \leqq 1$. Cf. Riesz, F., *Szeged Acta*, **5**, 23–54 (1930).

[3] Cf. Radon, J., *Wiener Berichte*, **122** (IIa), 1295–1438 (1913); Banach, S., *Théorie des Opérations Linéaires*, 1932; Kaczmarz, S., and Steinhaus, H., *Theorie der Orthogonalreihen*, 1935. The convergence theorems of Banach can be extended to denumerable B_*-systems.

[4] In other words, a complete linear metric space is also "hyper-complete." By a directed system we mean one in which there is a transitive relation with composition property. See Moore, these PROCEEDINGS, **1**, 628–632 (1915); Birkhoff, G., *Ann. Math.*, **38**, 39–56 (1937).

[5] *Fourier Transforms in the Complex Domain*, 1934, pp. 100–106.

[6] We make frequent use of results of two papers by Tseng in the *Science Reports of Tsing-Hua Univ.*, Series A: "On Schmidt's problem...," **3**, 299–316 (1936); "Expansions according to an arbitrary system...," forthcoming in Vol. 4, No. 4.

[7] Pell, *Trans. Amer. Math. Soc.*, **12**, 135–164 (1911).

[8] Lewin, *Math. Ztschr.*, **32**, 491–511 (1930).

[9] Cf. Remark (α) to Theorem C.

[10] Duffin, R. J., and Eachus, J. J., *Bull. Amer. Math. Soc.*, abstract no. 265, **46**, 415 (1940).

[11] Stone, M. H., *Linear Transformations in Hilbert Space*, 1932, p. 332.

REPRESENTATION OF ONE-PARAMETER SEMI-GROUPS OF LINEAR TRANSFORMATIONS

BY EINAR HILLE

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY

Communicated April 7, 1942

1. Let E be a separable normed complete linear vector space. Let T_s be defined for $s > 0$ as a linear bounded transformation on E to E such that

$$T_s T_t = T_t T_s = T_{s+t}, \; s > 0, t > 0. \tag{1.1}$$

We suppose further:

(1) T_s is weakly measurable for $s > 0$,
(2) $||T_s|| \leq 1$, $s > 0$,
(3) $T_s(E)$ is dense in E.

Here (2) can be replaced by $||T_s|| \leq M$ for $0 < s < 1$ without inconvenience but with some unessential modifications of the results. It is enough that (3) holds for a single s, it will then hold for all. If (3) is not satisfied in the

original space, we can often find a subspace where it holds and the discussion then applies to that subspace.[1]

On the basis of these assumptions we have proved strong continuity,[2] differentiability and representation theorems for semi-groups. The prototype of such results is the theorem of M. H. Stone on unitary groups in Hilbert space.[3] Our method is essentially that of Stone. Corresponding results for groups on spaces of type (B) have been announced by I. Gelfand and M. Fukamiya.[4]

We form

$$R(\lambda)x = -\int_0^\infty e^{-\lambda s} T_s x\, ds, \qquad x \in E,\ \Re(\lambda) > 0 \tag{1.2}$$

where the integral exists for instance in the sense of Bochner and Dunford.[5] This formula defines a linear transformation $R(\lambda)$ on E to E for $\Re(\lambda) > 0$ and $||R(\lambda)|| \leq [\Re(\lambda)]^{-1}$. The range $R(\lambda)(E)$ of $R(\lambda)$ is dense in E for every λ. Further $(\lambda - \mu)R(\lambda)R(\mu) = R(\lambda) - R(\mu)$ for $\Re(\lambda) > 0$, $\Re(\mu) > 0$ and $R(\lambda_0)x = 0$ implies $x = 0$.

$R(\lambda)$ is the resolvent of a closed linear transformation A whose domain of definition, $D(A)$, say, includes $R(\lambda)(E)$ for every λ and is, therefore, dense in E. We have

$$R(\lambda)(A - \lambda I) = I \text{ in } D(A),\ (A - \lambda I)R(\lambda) = I \text{ in } E, \tag{1.3}$$

and for every $x \in D(A)$ we have in the sense of strong convergence

$$Ax = \lim_{h \to 0} A_h x,\ A_h = \frac{1}{h}[T_h - I]. \tag{1.4}$$

The point spectrum of A may be dense in $\Re(\lambda) \leq 0$. From the fact that $D(A)$ is dense in E we conclude that for $h \to 0$

$$T_h x \to x,\ x \in E, \tag{1.5}$$

again in the sense of strong convergence.

2. With the aid of (1.2) we can derive a number of representations of the semi-group. The classical inversion formula of Laplace-Stieltjes integrals can be made to yield the result that for $c > 0$, $s \geq 0$ and $x \in E$

$$\int_0^s T_u x\, du = -\lim_{\omega \to \infty} (2\pi i)^{-1} \int_{c - i\omega}^{c + i\omega} e^{s\lambda} R(\lambda) x \frac{d\lambda}{\lambda}, \tag{2.1}$$

where the limit exists in the strong sense. If $x \in D(A)$, the integral obtained by letting $\omega \to \infty$ is convergent. Further, for $s > 0$, $x \in D(A)$

$$T_s x = -\lim_{\omega \to \infty} (2\pi i)^{-1} \int_{c - i\omega}^{c + i\omega} e^{s\lambda} R(\lambda) x\, d\lambda \tag{2.2}$$

in the strong sense. For $s = 0$, the formula gives $\frac{1}{2}x$ instead of x. If

the limit in (2.2) be replaced by the $(C, 1)$-limit, we can omit the restriction $x \in D(A)$.

The right-hand side of formula (2.2) can be interpreted as a definition of $\exp(sA)x$ in the sense of operational calculus,[6] so that

$$T_s x = \exp(sA)x, \; x \in D(A), \; s > 0. \tag{2.3}$$

Other interpretations of $\exp(sA)$ are obtainable by our methods. Thus

$$\int_0^s T_u x du = \lim_{h \to 0} \int_0^s \exp(uA_h)x du \tag{2.4}$$

for $s \geq 0$ and every $x \in E$, while

$$T_s x = \lim_{h \to 0} \exp(sA_h)x \tag{2.5}$$

for $x \in D(A)$. Here

$$\exp(sA_h)x = \sum_0^\infty \frac{s^n}{n!} A_h^n x. \tag{2.6}$$

The proof of these relations is obtained by observing that $\exp(sA_h)$ defines a semi-group for $s > 0$, the corresponding differential operator being A_h with a resolvent $R_h(\lambda)$, and the latter can be shown to converge uniformly to $R(\lambda)$, sufficiently regularly with respect to λ, so that if $R(\lambda)$ is replaced by $R_h(\lambda)$ in (2.1), we can pass to the limit with h under the sign of integration.[7]

For the case of a group I. Gelfand has proposed the interpretation

$$\exp(sA)x = \sum_0^\infty \frac{s^n}{n!} A^n x. \tag{2.7}$$

Gelfand's method appears to break down for semi-groups and it is not clear, in general, that there will exist elements belonging to all sets $D(A^n)$ for which the series is convergent in any sense whatsoever. For special semi-groups the situation may be different. Thus if T_s is defined on a Lebesgue space and T_s commutes with real translations on the point variable, then it is an easy matter to show that there exists a subspace E_0 dense on E such that the series in (2.7) converges for every $x \in E_0$ and every finite real or complex s in the sense that the sum of the norms of the terms is convergent.

[1] Condition (3) is used in proving that $R(\lambda)(E)$ and hence also $D(A)$ are dense in E. We also use the separability of the space at this point.

[2] Strong continuity for $s > 0$ but not for $s = 0$ has been proved by Dunford, N., "On One-Parameter Groups of Linear Transformations," *Ann. Math.*, ser. 2, **39**, 569–573 (1938), under more general assumptions.

[3] These PROCEEDINGS, **16**, 173–174 (1930), and *Ann. Math.*, ser. 2, **33**, 643–648 (1932).

[4] Gelfand, I., "On One-Parametrical Groups of Operators in a Normed Space," *Compt. Rend. Acad. Sci. U. R. S. S.*, **25**, 713–718 (1939); Fukamiya, M., "On One-Parameter Groups of Operators," *Proc. Imp. Acad. Tokyo*, **16**, 262–265 (1940). Fukamiya assumes separability; his proof, partly somewhat fragmentary, uses the method of

Stone, but does not go beyond the existence of A and $R(\lambda)$ and stops short of representation theorems. Gelfand in part uses an unpublished notion of integration of his own; he claims to be able to dispense with separability, and arrives at the representation (2.7) below, valid in a subspace dense in E.

[5] Bochner, S., "Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind," *Fund. Math.*, **20**, 262–276 (1933); Dunford, N., "Integration in General Analysis," *Trans. Amer. Math. Soc.*, **37**, 441–453 (1935). Separability plus weak measurability implies that the integrand is measurable in the sense of Bochner (see Pettis, B. J., "Integration in Vector Spaces," *Trans. Amer. Math. Soc.*, **44**, 277–304 (1938), (Theorem 1.1 and Corollary 1.11) and the norm is measurable and dominated by an integrable function.

[6] Cf. Stone, M. H., *Linear Transformations in Hilbert Space*, New York, 1932, Chapter VI.

[7] The author has found a simpler proof of (2.5) which is valid for all x and E—Condition (3) can be replaced by the weaker condition (3*): The least linear hull of the range spaces is dense in E. This condition is necessary for the validity of (1.5). [Added in proof.]

ON A CLASS OF SURFACES

BY G. FUBINI

THE INSTITUTE FOR ADVANCED STUDY, PRINCETON, N. J.

Communicated March 30, 1942

An interesting problem of the differential projective geometry is the problem of finding all the surfaces S which possess at least two systems of curves such that the curves of the same system are projective to each other. Many particular surfaces S are well known (see below). We can suppose that the two systems of curves on S are given by the equations $u = \text{const.}$ and $v = \text{const.}$ Let us suppose that the line $\Gamma_{v=0}$ defined by $v = 0$ is given by the equation

$$x_i = b_i(u) \qquad (i = 1, 2, 3, 4) \tag{1}$$

in which the b are functions of u, and the x are projective homogeneous coordinates, and that the line $\Gamma_{u=0}$ is given by

$$x_i = e_i(v) \qquad (e_i \text{ functions of } v). \tag{2}$$

The point $A(u = v = 0)$ will be the point

$$x_i = \lambda_i = b_i(0) = e_i(0). \tag{3}$$

If the lines $v = \text{const.}$ are projective to each other, every line $v = \text{const.}$ can be deduced from the line $\Gamma_{v=0}$ by means of a projectivity T_v, the coefficients a_{ik} of which $(i, k = 1, 2, 3, 4)$ are functions of v such that the determinant of the a_{ik} is different from zero and that, for $v = 0$, the pro-

jectivity T_v is the identity. Therefore the coördinates x of the most general point of the surface S will be given by

$$x_i = \sum_k a_{ik}(v)b_k(U) \tag{4}$$

where

$$U = U(u, v) \tag{5}$$

is a function of u, v such that for $v = 0$, $U = u$. The point $u = u_0$ of the line $\Gamma_{v=0}$ is carried by the projectivity T_v into the point

$$x_i = \sum_k a_{ik}(v)b_k(u_0). \tag{6}$$

By comparing (4) and (6) we find that

$$U(u, v) = u_0 = \text{const.} \tag{7}$$

is the equation of the *trajectories* of the projectivities T_v.

In the same manner, by considering the lines $u = \text{const.}$, we shall find another system of projectivities T_u, the coefficients c_{ik} of which are functions of u; their determinant is different from zero; the projectivity T_u, for $u = 0$, is the identity. Therefore every point of S can be defined also by the equation

$$x_i = \sum_k c_{ik}(u)e_k(V) \quad (V = V(u, v))(V = v \text{ for } u = 0) \quad (4)_{\text{bis}}$$

in which V is a function of u, v, and the equation $V = \text{const.}$ determines the trajectories of the projectivities T_u. Since (4) and $(4)_{\text{bis}}$ determine the same point, and the x are homogeneous coördinates, we obtain the following system of equations

$$\sum_k a_{ik}(v)b_k(U) = R(u, v) \sum_h c_{ih}(u)e_h(V) \quad (i = 1, 2, 3, 4) \quad \text{(I)}$$

in which R is a function of u, v finite and different from zero. It seems very difficult to find the most general solution of (I); and the problem is also difficult even if we suppose that the surface S possesses a third system of curves projective to each other. Particular solutions are well known: the quadrics, the surfaces of rotation, the surfaces of translation, the ruled surfaces, the asymptotic lines of which are curves belonging to a linear complex, etc. We know that it often happens that the trajectories $U = \text{const.}$ and $V = \text{const.}$ are identical with the lines $u = \text{const.}$, $v = \text{const.}$, so that equation (I) becomes (by supposing $U = u$, $V = v$)

$$\sum_k a_{ik}(v)b_k(u) = R(u, v) \sum_h c_{ih}(u)e_h(v). \tag{II}$$

This note is devoted to the solution of (II) under the hypothesis that neither the line $v = 0$ nor the line $u = 0$ belongs to an algebraic surface. Therefore neither the lines $v =$ const. nor the lines $u =$ const. will belong to an algebraic surface. (It follows that they cannot be straight lines or plane curves.) For this exceptional case other methods are necessary.

According to our hypotheses we can find two constants, $\alpha \neq 0$, $\gamma \neq 0$, such that

$$a_{ik}(0) = \alpha\delta_{ik} \quad c_{ik}(0) = \gamma\delta_{ik} \quad (\delta_{ii} = 1;\ \delta_{ik} = 0 \text{ if } i \neq k).$$

Equation (II) is equivalent to the equation

$$\sum_k \bar{a}_{ik}(v)\bar{b}_k(u) = \overline{R}(u, v) \sum_h \bar{c}_{ih}(u)\bar{e}_h(v)$$

in which

$$\bar{a}_{ik} = \frac{a_{ik}}{\alpha}, \qquad \bar{b}_k(u) = \alpha \frac{b_k(u)}{R(u, 0)} \qquad \overline{R}(u, v) = \frac{R(u, v)R(0, 0)}{R(u, 0)R(0, v)}$$

$$\bar{c}_{ik} = \frac{c_{ik}}{\gamma} \qquad \bar{e}_h(v) = \gamma \frac{R(0, v)}{R(0, 0)} e_h(v).$$

Since the coördinates are homogeneous, the point $x_k = \bar{b}_k(u)$ is identical with the point $x_k = b_k$ and the point $x_k = \bar{e}_k$ is identical with the point $x_k = b_k$. We find that

$$\bar{a}_{ik}(0) = \bar{c}_{ik}(0) = \delta_{ik} \qquad \overline{R}(0, v) = \overline{R}(u, 0) = \overline{R}(0, 0) = 1.$$

By changing notation, and disregarding the dashes, we obtain the equations

$$\frac{\sum_k a_{ik}(v)b_k(u) = R(u, v) \sum_h c_{ih}(u)e_h(v) \qquad (i = 1, 2, 3, 4)}{a_{ik}(0) = c_{ik}(0) = \delta_{ik} \qquad R(0, v) = R(u, 0) = R(0, 0) = 1} \tag{8}$$

and by supposing $u = v = 0$, we obtain again $b_i(0) = e_i(0)$, so that we can always write equation (3) with new values of the λ. By supposing only $u = 0$, or $v = 0$, we get

$$\begin{aligned} b_i(u) &= \sum_h \lambda_h c_{ih}(u), \\ e_i(v) &= \sum_k \lambda_k a_{ik}(v). \end{aligned} \tag{9}$$

By means of a projective change of coördinates we could suppose that the point $u = v = 0$ (the coördinates of which are the λ_i) is the point (0, 0, 0, 1). In this case we might write

$$b_i(u) = c_{i4}(u) \qquad e_i(v) = a_{i4}(v).$$

But this simplification is of no importance for what follows. Let us choose

four values of u ($u = u_1, u_2, u_3, u_4$) and let us consider the determinant Δ_1 the rth row of which ($r = 1, 2, 3, 4$) is

$$e_r(v) \qquad \sum_h c_{rh}(u_2)e_h(v) \qquad \sum_h c_{rh}(u_3)e_h(v) \qquad \sum_h c_{rh}(u_4)e_h(v).$$

From (8) and (9) we deduce that it is equal to the determinant, the rth row of which is

$$\sum_k \lambda_k a_{rk}(v) \qquad \frac{1}{R(u_2, v)}\sum_k a_{rk}(v)b_k(u_2) \ldots \frac{1}{R(u_4, v)}\sum_k a_{rk}(v)b_k(u_4).$$

Therefore the determinant Δ_1 is equal to

$$\Delta_1 = \frac{AB_1}{R(u_2, v)R(u_3, v)R(u_4, v)} \tag{10}$$

in which A is the determinant of the α_{ik} and

$$B_1 = \begin{vmatrix} \lambda_1 & b_1(u_2) & b_1(u_3) & b_1(u_4) \\ \lambda_2 & b_2(u_2) & b_2(u_3) & b_2(u_4) \\ \lambda_3 & b_3(u_2) & b_3(u_3) & b_3(u_4) \\ \lambda_4 & b_4(u_2) & b_4(u_3) & b_4(u_4) \end{vmatrix}. \tag{11}$$

We recall that $\lambda_i = b_i(0)$. Since the curve $\Gamma_{v=0}$ is not plane, we can suppose that $B_1 \neq 0$. (We already know that A and $\frac{1}{R(u, v)}$ are different from zero, too.)

In the same manner one finds that the determinant Δ, the rth row of which is

$$\sum_h c_{rh}(u_1)e_h(v) \qquad \sum c_{rh}(u_2)e_h(v) \qquad \sum c_{rh}(u_3)e_h(v) \qquad \sum c_{rh}(u_4)e_h(v),$$

is equal to

$$\Delta = \frac{AB}{R(u_1, v)R(u_2, v)R(u_3, v)R(u_4, v)} \tag{12}$$

in which B is the determinant of the $b_r(u_s)$,

$$B = \begin{vmatrix} b_1(u_1) \ldots\ldots b_1(u_4) \\ \ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots \\ b_4(u_1) \ldots\ldots b_4(u_4) \end{vmatrix}. \tag{13}$$

From these identities we deduce that

$$\frac{\Delta_1}{\Delta} = \frac{B_1}{B}R(u_1, v) \text{ or } R(u_1, v) = \frac{\Delta_1 : B_1}{\Delta : B}$$

Let us now consider u_2, u_3, u_4 as constants, and $u_1 = u$ as a variable. We realize without difficulty that Δ_1 is a homogeneous polynomial $P(e)$ in e_i of fourth degree, the coefficients of which are functions of *only* u_2, u_3, u_4 and are therefore constants. Also B_1 is constant. On the other hand, $\Delta:B$ is a homogeneous polynomial $P_u(e)$ of fourth degree in e_i, the coefficients of which are functions of u. And we have seen that both of these polynomials are different from zero. Therefore

$$R(u, v) = \frac{P(e)}{P_u(e)}. \tag{14}$$

We could arrive at this result in another way. Let us write $b_{kr} = b_k(u_r)$ and indicate by β_{kr} the algebraic complement of b_{kr} in the determinant B of the $b_k(u_r)$.

By supposing $u = u_1, u_2, u_3, u_4$ in (8) we obtain four linear equations in the four unknowns $a_{i1}, a_{i2}, a_{i3}, a_{i4}$; by solving these equations in the usual way, we obtain

$$a_{ij} = \sum_r \beta_{jr} R(u_r, v) \sum_h c_{ih}(u_r) e_h(v)$$

and from (9)

$$e_i(v) = \sum_{j,r} \lambda_j \beta_{jr} R(u_r, v) \sum_h c_{ih}(u_r) e_h(v) \qquad (i = 1, 2, 3, 4).$$

I now solve these four equations by considering the

$$z_r = \sum_j \lambda_j \beta_{jr} R(u_r, v) \qquad (r = 1, 2, 3, 4)$$

as unknowns. I find (by supposing $r = 1$) that the value of

$$\sum_j \lambda_j \beta_{j1} R(u_1, v) = \frac{B_1}{B} R(u_1, v)$$

is equal to $\Delta_1:\Delta$.

In the same way, by interchanging the parameters u, v, and therefore by interchanging R and $1/R$, I shall find that

$$\frac{1}{R(u, v)} = \frac{Q(b)}{Q_v(b)} \tag{15}$$

in which Q and Q_v are homogeneous polynomials of fourth degree in b_i: the coefficients of Q are constants, the coefficients of Q_v are functions of v. By comparing (14) and (15) we obtain

$$P_u(e) = \frac{P(e)Q(b)}{Q_v(b)} \text{ or } P_u(e)Q_v(b) = P(e)Q(b). \tag{16}$$

This equation is an identity in u, v if e_i and b_i are the coördinates of a point whatsoever of the curves $\Gamma_{u=0}$ and $\Gamma_{v=0}$. If I write, for instance, $x_i = e_i(v)$, I cannot state that this equation is an identity in x_1, x_2, x_3, x_4 (for every value of u, because $Q_v(b)$ is a function of v). But it is very easy to consider this equation from a deeper point of view. Let ϵ be the coefficients of $P_u(e)$, which are functions of u; and let us write

$$P_u(e) = \epsilon_{1111}e_1^4 + \epsilon_{1112}e_1^3e_2 + \ldots + \epsilon_{1234}e_1e_2e_3e_4 + \ldots + \epsilon_{4444}e_4^4 = \frac{P(e)Q(b)}{Q_v(b)}. \tag{17}$$

I consider now 35 values of v, for instance $v = v_r$ $(r = 1, 2, \ldots, 35)$. By supposing $v = v_r$ in (17), I obtain 35 linear equations in the 35 coefficients and, if the curve $\Gamma_{u=0}$ $(x_i = e_i(v))$ does not belong to an algebraic surface of degree $m \leq 4$ I can suppose that the v_r are such that the determinant of the coefficients of the ϵ in these 35 equations is different from zero. We can therefore solve these equations with respect to the ϵ; and, by remarking that, for $v = v_r$, the last member of (17) is equal to $Q(b)/K_r(b)$ in which K_r is a new homogeneous polynomial of fourth degree in the b, with constant coefficients, I deduce that every ϵ is equal to an expression

$$\sum h_r \frac{Q(b)}{K_r(b)} \qquad (h_r = \text{const.}).$$

(The constants h_r change if we change the coefficient ϵ which we wish to calculate.) From (17) we deduce consequently that

$$P_u(e) = \sum \pi_r(e) \frac{Q(b)}{K_r(b)} \tag{18}$$

in which $\pi_r(e)$ is a homogeneous polynomial of degree 4 in e_i with constant coefficients. In the same manner we prove that

$$Q_v(b) = \sum T_S(b) \frac{P(e)}{H_S(e)} \tag{19}$$

in which T_S and H_S are homogeneous polynomials of fourth degree (in the b or in the e) with constant coefficients. By substituting (18) and (19) in (16), we obtain

$$\sum \frac{\pi_r(e)}{K_r(b)} \sum \frac{T_s(b)}{H_s(b)} = 1. \tag{20}$$

For every value of v this equation must be an identity in the b; if it were not so, equation (2) would be an algebraic equation in the b, and the curve $x = b_i(u)$ would belong to an algebraic surface; which is contrary to our

hypothesis. Let us give to v a constant value, such that all the $H_s(e)$ are different from zero. The equation

$$\sum \frac{T_s(b)}{H_s(e)} = 0 \tag{21}$$

defines a surface Σ, if the b are considered as homogeneous coördinates of a point. And, since (20) is an identity in the b (for every value of v), it is obviously necessary that the points of this surface Σ satisfy the equation

$$0 = K_1(b) \qquad K_2(b) \ldots \tag{22}$$

$\left(\text{The second member is the product of the denominators of } \frac{\pi_r(e)}{K_r(b)}.\right)$ But (22) does not depend on the value of v, whereas (21) depends on this value. This is possible only if

$$\sum \frac{T_s(b)}{H_s(e)} = \psi(v)L(b) \tag{23}$$

in which ψ is a function only of v and L is a homogeneous polynomial of fourth degree in b with constant coefficients. In the same manner I find that

$$\sum \frac{\pi_r(e)}{K_r(b)} = \psi(u)M(e) \tag{24}$$

in which $\psi(u)$ is a function of only u and M is a homogeneous polynomial of fourth degree in e with constant coefficients. Equation (20) becomes

$$\varphi(v)M(e) = \frac{1}{\psi(u)L(v)}.$$

Since the first member is independent of u, and the second of v, both are equal to the same constant C

$$\pi(v) = \frac{C}{M(e)}; \quad \psi(u) = \frac{1}{CL(b)}. \tag{25}$$

From (18), (24), (25), and from (14), we obtain

$$P_u(e) = Q(b)\frac{M(e)}{CL(b)}, \quad R(u, v) = \frac{P(e)}{M(t)}\frac{\Lambda(b)}{Q(b)} \quad [\Lambda(b) = CL(b)] \tag{26}$$

in which all the polynomials are homogeneous, of fourth degree with constant coefficients. But $R(0, v) = 1$, and $\Lambda(b)/Q(b)$ is a constant, if $u = 0$. Therefore $P(e)/M(e) = k$ const., and this equation must be an identity, since the curve $\Gamma_{u=0}$ $(x_i = e_i(v))$ does not belong to an algebraic surface. In the same manner one realizes that $\Lambda(b)/M(b)$ is a constant. From (26)

we deduce consequently that R is a constant. And, since $R(0, v) = R(u, 0) = R(0, 0) = 1$, we deduce that

$$R(u, v) = 1.$$

Equation (8) becomes

$$\sum a_{ik}(v)b_k(u) = \sum c_{ih}(v)e_h(v).$$

By giving to u four particular values $u = u_1, u_2, u_3, u_4$, and solving these equations with respect to $a_{i1}, a_{i2}, a_{i3}, a_{i4}$, we obtain

$$a_{ij} = \sum_h \lambda_{ijh}e_h(v) \qquad (\lambda_{ijh} = \text{const.})$$

and therefore

$$\sum_{k,\,h} \lambda_{ikh}b_k(u)e_h(v) = \sum c_{ir}(u)e_r(v).$$

As usual, this equation is an identity, and therefore

$$C_{ir} = \sum_k \lambda_{ikr}b_k(u).$$

The parametric equations of our surfaces are

$$x_i = \sum_{k,\,h} \lambda_{ikh}b_k(u)e_h(v)$$

with λ arbitrary constants.

The existence of these surfaces S was obvious *a priori*. It may perhaps be interesting to remark that, besides these surfaces, there is no other surface S, if neither the curve $u = 0$, nor the curve $v = 0$ belongs to an algebraic surface. If it is not so, for instance, if S is a ruled surface and the line $u = 0$ is a generatrix, the theorem can be completely false.

GALACTIC AND EXTRAGALACTIC STUDIES, XIII. NOTE ON THE COMPARATIVE DIAMETERS OF SPHEROIDAL AND SPIRAL GALAXIES

BY HARLOW SHAPLEY

HARVARD COLLEGE OBSERVATORY

Communicated April 13, 1942

From the photometric results presented in this communication it appears that contrary to some earlier indications the open spirals, with their supergiant stars and superficial structure of spiral arms, are not appreciably larger than spheroidal galaxies, which seem to be devoid of highly luminous stars and free of distributional irregularities. It follows that if in the course of time spiral arms appear in a flattened spheroidal system, they probably should not be treated as an extension or an expansion outward from the nuclear part of the galaxy, but as a development of structure well within the main body of the system. Or if, with the direction of evolution reversed, it be assumed that the spiral arms, supergiant stars and diffuse nebulosities of the spiral systems can eventually disappear into the structureless smooth form of the typical spheroidal galaxy, then that transformation also should be treated as an internal readjustment, and not as a contraction; for at any given considerable distance from the nucleus the amount of light (and probably of mass) is now found to be about the same for spheroidal and spiral systems.* Apparently there is no important "growth" in dimensions along the sequence of forms.†

The spiral arms as observed are a phenomenon of only the inner half of an average galaxy. Moreover, less than twenty per cent of the light of a spiral galaxy is, on the average, in its spiral arms. The remainder is mostly in the commonly overlooked background in which the spires are embedded.

Since galaxies, like planetary atmospheres, probably fade out indefinitely, the extreme diameters are not very important in themselves (if the peripheral masses are negligible); but the dimensions out to a given light or mass density become very significant if evolutionary trends are under consideration and comparative sizes and gradients must be discussed.

1. A densitometric analysis of photographs of 123 bright northern galaxies was recently undertaken at the Harvard Observatory, chiefly to determine the density gradients and thereby to contribute to our fragmentary knowledge of the distribution of light and mass in large stellar systems of various types. Dealing comparatively with a large number of objects of all types, this work supplements the detailed analyses for individual galaxies by Reynolds, Hubble, Oort, Shirley and Redmond, Lind-

blad, Seyfert and others. The treatment of the gradients by Dr. Shirley Patterson will soon be published in the *Harvard Annals*.

A by-product of Miss Patterson's photometry has been the measurement of boundaries. From this material the quantitative intercomparison of the dimensions of spheroidal and spiral galaxies now becomes possible. The observational material is not extensive, especially for the spheroidal type; but it seems to be considerably the best now available for this particular problem, and is, in fact, about the first that seems to be precisely suited to the intercomparison of dimensions.

Although it is clear that the interpretation of the well-known sequence of types requires quantitative information of the sort now obtained on the relative dimensions, we should not overvalue the results for dynamical analysis, because our photographs measure only the distribution of light; the mass distribution may be otherwise. The ratio of mass to luminosity may, indeed, vary from type to type; almost certainly it varies from center to periphery in some of the galaxies.

2. From a general inspection of photographs, the conclusion has been easily reached in the past that spheroidal galaxies are on the average much smaller than the spirals, and that the more open spirals are largest of all. For instance, Hubble has tabulated the following results from reflector plates available at Mount Wilson:[1]

TYPE	DIAMETER	TYPE	DIAMETER
E0	0.58 kpc	Sa	1.84 kpc
E3	0.86	Sb	2.33
E7	1.47	Sc	2.92

He recognizes that these diameters refer to the "main bodies" of the galaxies. They show a marked and significant trend, with the average spiral having more than twice the linear diameter of the average spheroidal galaxy. In fact, the spherical E0 type, as measured by Hubble, compares more closely with the nucleus or inner disc of type Sc than with the five-times-larger spiral itself, and this circumstance has naturally suggested that spiral arms are dynamical expulsions or extensions from the nucleus, if there be assumed a developmental sequence in the direction E0-E7-Sa-Sc.

This same trend in the measured dimensions is evident in the Shapley-Ames catalog, where the diameters are also based on direct eyepiece measurement of reflector or refractor plates;[2] but when the diameters of the galaxies were more closely studied on a homogeneous series of long exposure Harvard plates,[3] the "growth" became less pronounced and the measured linear diameters of all galaxies were greatly increased over the previously used values. For example, a comparison of these Harvard measures of 1934 with Mount Wilson results shows that for forty-two

TABLE 1

DIAMETERS OF 112 EXTERNAL GALAXIES

NGC	FSP TYPE	PG. MAG.	DIAM. IN kpc	NGC	FSP TYPE	PG. MAG.	DIAM. IN kpc
450	Sc2	12.6	5.0	4429	Sa	11.3	5.3
470	Sc2	12.7	3.6	4435	S0	11.7	4.4
474	E2.3	13.2	4.0	4442	SBa	11.5	4.1
488	Sa	11.8	4.4	4450	Sc1	11.1	4.7
628	Sc1	11.8	5.6	4459	E2.4	11.5	3.6
991	Sc2	12.7	5.4	4472	E1.0	10.1	7.0
1022	Sc2	12.0	2.7	4473	E4.1	11.4	4.6
1035	Sc1:	12.8	4.7	4486	E1.4	10.2	4.5
1042	Sc1	12.5	6.0	4490	Sc3	10.8	2.4
1052	E2.8	11.6	3.1	4501	Sc2	10.6	2.8
1084	Sc3	11.2	2.3	4503	S0	12.8	2.9
2403	Sc1	10.2	4.4	4517	(Sc)	11.6	8.4
2903	Sb1	10.3	4.0	4526	Sa	10.7	5.2
3031	Sb1	8.1	4.2	4527	Sc1	11.3	4.9
3115	E6.2	9.8	2.0	4535	Sc1	11.0	5.3
3184	Sc2	11.8	3.0	4536	Sc2	11.2	3.4
3198	Sc2	11.7	4.5	4548	Sc1	11.9	5.1
3338	Sc2	12.2	5.0	4550	Sa	12.8	2.9
3351	Sc1	11.5	5.4	4552	E0.0	11.7	4.8
3368	Sa	10.4	2.9	4559	Sc1	10.7	3.8
3556	Sc2	11.0	3.9	4564	S0	12.1	3.1
3623	Sc2	10.5	3.2	4565	Sb1	10.7	5.8
3627	Sc2	9.9	2.8	4567	Sc2	12.3	3.4
3628	(Sb)	11.3	6.7	4568	Sc2	12.2	4.0
3631	Sc3	11.8	2.2	4569	Sc1	11.3	13.6
3917	Sc3	12.8	5.0	4571	Sc1	12.8	3.8
3953	Sc2	11.5	3.2	4579	Sc1	10.8	11.5
3992	Sc2	11.2	4.2	4621	E2.7	11.8	4.2
4013	Sc1	12.7:	6.1	4631	Sc1	9.6	6.3
4036	Sa	11.9	4.0	4647	Sc2	12.7	4.9
4051	Sc1	11.7	4.4	4649	E1.2	11.2	5.7
4088	Sc3	11.2	2.3	4654	Sc2	11.7	4.7
4096	Sc1	12.2	6.0	4656	Sc2	11.3	3.8
4100	Sc2	11.9	3.2	4725	Sb1	10.8	4.8
4111	S0	11.6	3.9	4736	Sb1	9.0	1.9
4138	Sa	12.2	4.2	4826	Sc2	8.0:	2.8
4143	Sa	12.2	2.6	5005	Sc3	11.3	2.7
4144	Sc2	12.4	7.2	5033	Sb1	11.8	6.5
4157	Sc1	12.0	5.1	5055	Sb1	10.5	5.6
4192	Sc1	11.1	7.2	5194	Sc1	10.1	3.3
4214	Sc1	10.7	4.1	5371	Sc2	11.7	2.6
4216	Sc1	11.2	6.9	5377	Sa	12.8	5.9
4236	Sc1	11.3:	6.2	5457	Sc2	9.0	2.9
4242	Sc2	11.8	4.3	5474	Sc2	11.7	2.7
4244	Sc1	11.0	8.9	5631	Sa	12.5	3.1
4254	Sc2	10.5	4.5	5660	Sc3	12.3	2.7
4258	Sb1	10.2	6.8	5676	Sc3	11.9	3.4
4274	Sa	11.7	4.1	5689	SBa	12.6	3.5
4281	Sa	12.6	4.4	5866	Sa	11.5	4.5
4303	Sc1	10.4	5.8	5907	Sc	11.8	8.8
4321	Sc1	10.5	6.3	5985	Sc1	12.2	4.1
4371	SBa	12.1	4.6	6015	Sc3	12.1	3.2
4374	E2.1	10.7	5.7	6503	Sc3	11.4	4.2
4382	E2.7	10.3	7.0	6946	Sc1	11.1:	4.7
4406	E1.0	10.9	7.3	7217	Sc1	12.2	2.9
4414	Sb1	11.3	4.2	7479	Sc2	12.3	3.8

spheroidal galaxies the diameters are increased by an average of 3.2 times, but for eighty spirals only 1.6 times. Thus the large difference between

the two general classes was lessened, even when only visual estimates were employed on non-standardized photographs.

At the same time some preliminary measurements with a microdensitometer of a few of the same series of Harvard plates (three-hour exposures with the Bruce refractor) indicated that when using simple eyepiece measures we are far from exhausting the evidence on dimensions. Since some galaxies are common to the various dimensional surveys, we can show the advantage of densitometric measurement by determining mean ratios of the measures of diameters by the various methods.[3] We have the ratios

Microdensitometer to Mount Wilson estimates = 4.1 (9 objects)
Microdensitometer to Harvard 1934 estimates = 1.6 (14 objects)

Since this early densitometric work was based on photographs that were not precisely standardized, a study of the overall dimensions of galax-

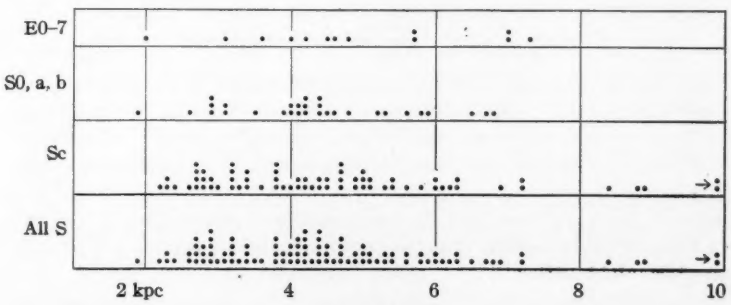

FIGURE 1

Distribution of diameters for 112 galaxies. Ordinates are numbers; abscissae are major diameters in kiloparsecs.

ies was then not carried further. The problem is now taken up again because suitable photographs have been made. The plates were exposed and developed under standard conditions by Miss Patterson. The 12-inch Metcalf refractor at Oak Ridge is especially suited to the work because of its large flat field and its suitable focal length. The plates have been systematically calibrated with a tube photometer, developed in a way that avoids Eberhard effect, and measured with a uniform technique. A description of the plate material and the methods involved will be described elsewhere.[4]

3. The linear diameters are given for 112 galaxies in table 1. Although they, too, have been analyzed photometrically, the nearby systems M 31,

32, 33, are omitted, as are also five irregular galaxies, and three classed as S without subdivision. The tabulated quantities indicate the extreme lengths of the measurable major axes. On a densitometer tracing the border of a galaxy is indicated by the clear rising of the density curve from the background of sky and plate fog. The important point in the present intercomparison is the strict comparability of the light intensity, for all types of galaxies, near the ends of the respective major axes.

The apparent magnitudes in table 1 are photoelectric (Stebbins and Whitford), when available, or photographic (*Harv. Ann.*, **88,** No. 2), except that 8.1 is now adopted for NGC 3031. The linear diameters are based on the best values of the distances at present available.

4. In figure 1 the distribution of the diameters is graphically shown for the various subtypes. The smallest major diameters are about 2000 parsecs; the median and mean values for all 112 objects are more than double that value, and are practically the same for all types represented in the figure. The averages can be summarized as follows:

TYPE	NUMBER	MEDIAN DIAMETER, KPC	MEAN DIAMETER, KPC	MEAN ERROR
E0–E7	13	4.6	4.88	±0.44
S0, a, b	30	4.2	4.35	±0.23
Sc	69	4.3	4.67	±0.25
All S	99	4.2	4.57	

In view of the dispersions and the uncertainties of measurement, we infer from these tabulated results that the galaxies of various types are of strictly comparable dimensions.

5. It could be argued that the large value of the average diameter of the spheroidal galaxies arises from the selection for densitometric work of the brightest and largest of that class. From various earlier studies it is known that the spheroidal and spiral systems are more or less evenly scattered throughout all observed magnitudes in clusters of galaxies as well as in open metagalactic space. In consequence the dispersions in absolute magnitude are much alike.

To test for evidence of selection in the densitometric study, let us consider all galaxies listed in the fairly homogeneous Shapley-Ames catalog between right ascensions $12^h.0$ and $13^h.0$ and declinations 0° and +20°, thus including practically all the recognized members of the Virgo cluster. As shown below, the mean apparent magnitude of the thirty-six spheroidal galaxies is 11.98; for the eighty spirals it is 12.09. The magnitude difference of 0.11 should be compared with the difference of 0.51 in the last column of the following tabulation where only those Virgo galaxies are included that were analyzed with the densitometer.

TYPE IN *Harv. Ann.*, VOL. 88, NO. 2	ALL IN VIRGO REGION	ANALYZED BY DENSITOMETER
E0–E7	$11^m.98$ (36)	$11^m.07$ (11)
S, S0, a, b, c	12.09 (80)	11.58 (27)
Magnitude difference	0.11	0.51

The tabulated quantities are mean magnitudes, with numbers of galaxies in parentheses.

Obviously there has been some selection of the bright spheroidal systems for the densitometric work—a selection on the basis of apparent (and absolute) magnitude, and it amounts in the mean to $\Delta m = 0^m.4$. To correct the average diameter, D, for this selection we have

$$\Delta \log D = -0.2\,\Delta m = -0.08$$

and to the mean and median diameters of the measured spheroidal galaxies should be applied the reduction factor 0.83. The results corrected for selection are, therefore,

Spheroidal: median diameter 3.8 kpc, mean diameter 4.05 kpc
Spiral: median diameter 4.2 kpc, mean diameter 4.57 kpc

This reduction still leaves spheroidal galaxies much larger than the part of the spirals occupied by recognizable spiral arms, and leaves us with the conclusion, mentioned in the introduction, that the two main types of external galaxies are of essentially the same linear dimensions.

* It is, of course, not strictly necessary to assume that galaxies actually develop from one type to another. But a static universe (so far as galaxy evolution is concerned), or a very recent "creation," or a common unknown parent of all galaxy forms, are among the less comfortable of possible hypotheses.

† These conclusions are not incompatible with Hubble's recent important observations on the nature of the transition form S0 and on the first appearance (or vanishing) of disc and spiral structure in the spheroidal-spiral series. See also Randers, *Mt. Wilson Contr.*, No. 634, p. 282 (1940).

[1] *The Realm of the Nebulae*, p. 178 (1936).

[2] *Harv. Ann.*, **88**, No. 2 (1932).

[3] *Ibid.*, **88**, No. 4 (1934); also *Harv. Bull.*, 895, p. 22 ff. (1934).

[4] *Harv. Ann.*, **88**, No. 7 (in press).

GALACTIC AND EXTRAGALACTIC STUDIES, XIV. ON THE MAGNITUDE DISPERSION IN THE PERIOD-LUMINOSITY RELATION

BY HARLOW SHAPLEY, VIRGINIA MCKIBBEN AND RICHARD A. CRAIG

HARVARD COLLEGE OBSERVATORY

Communicated April 13, 1942

The number of Cepheid variable stars in the Small Magellanic Cloud for which we now have periods and light curves has been increased to 564. The data on the dispersion of median magnitudes about the mean period-luminosity curve are therefore sufficiently abundant to merit preliminary examination. Although it is still impossible to disentangle the various causes of deviation, we are able to study the effect of doubling (superposed images), to reëxamine the evidence for space absorption within the Cloud, and to test the fit of the adopted period-luminosity curve.

The new periods—about two hundred and fifty in number—which have been determined during the past year, will be published soon in a Circular of the Harvard College Observatory.

1. Causes of Dispersion.—The spread of the median magnitudes for the Cepheids in any compact stellar system can be attributed to some or, more probably, to all of the following six principal factors:

(1) irregular space absorption in the system,
(2) superposition of star images,
(3) extent of the system in the line of sight,
(4) inherent spread of luminosities for stars of the same period,
(5) Eberhard effect and background haze,
(6) observational and computational uncertainties, including magnitude-sequence irregularity.

The first factor throws the point for the affected star below (fainter than) the average period-luminosity curve; the second throws it above, and somewhat diminishes the range. The other four can work in both directions.

2. The Period-Luminosity Curve.—The present status of the relation of apparent photographic median magnitude to the logarithm of the period is exhibited in the usual form in figure 1, where the curve is that adopted earlier.[1] The curve was based on data for 307 stars but reasonably fits the 564 now plotted. The survey for variables with median magnitudes fainter than 16.8 is probably not complete because of faintness and crowding; the lower part of the curve was drawn with this probable deficiency in mind.

3. The Scatter Diagram of Magnitude Residuals.—The dispersion in the magnitudes is best shown by the scatter diagram in figure 2, where the

abscissae are logarithms of the period and the ordinates are deviations in terms of magnitude. The errors in the periods are relatively so small that we ignore them in this discussion and treat the period-luminosity curve as a regression relation. The factors listed in the first section above produce significant dispersions in magnitude only, except that number (4), inherent spread, could of course be associated with length of period as well as to luminosity.

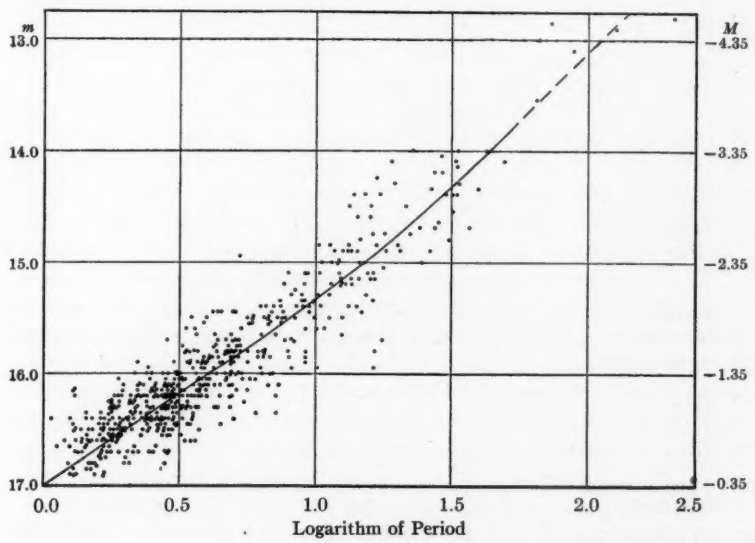

FIGURE 1

Period-luminosity observations for Small Magellanic Cloud. Ordinates are apparent magnitudes (left) and absolute magnitudes (right); abscissae are logarithms of the periods in days.

The plotted residuals in figure 2 have been computed for each variable from the relation

$$\Delta m = \dot{m}' - \dot{m}$$

where $\dot{m}'$ is the observed median photographic magnitude and

$$\dot{m} = 17.07 - 1.74 \log P$$

This last relation is the straight line found, in the earlier discussion,[1] to represent the period-luminosity relation most satisfactorily for apparent median magnitudes in the Small Cloud. The adopted non-linear period-luminosity curve deviates only slightly, as shown in figure 4, from the

straight line; and in examining the relative dispersions for different periods, luminosities and positions in the Cloud, the linear relation suffices.

The spread in the median magnitudes is much the same throughout the whole range of periods (figures 2 and 4). At about sixteen days it is larger than average, but this is chiefly due to a few points far below the curve for which some of the magnitudes appear to suffer through intra-Cloud

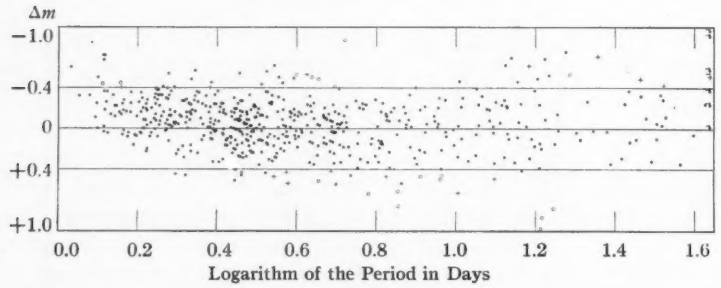

FIGURE 2

Magnitude residuals from the linear period-luminosity relation. For residuals greater than ±0.4, open circles refer to variables in the nucleus, crosses to variables in border regions.

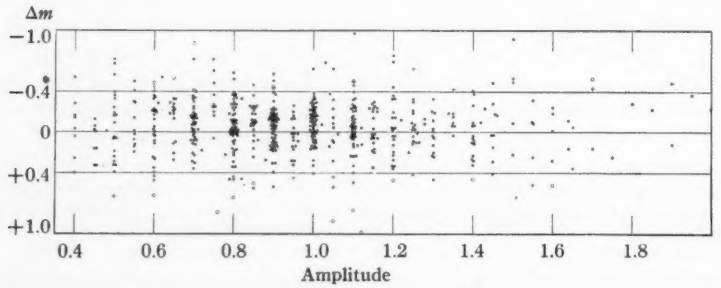

FIGURE 3

Magnitude residuals from the linear period-luminosity curve against amplitudes of variation. Open circles and crosses as for figure 2.

absorption; and it is somewhat less for periods of 1.5 days, almost certainly because of the incompleteness of our records for faint magnitudes.

4. Dispersion and Amplitude.—It is possible that the points lying well above the average curve (figures 1 and 2) refer in part to stars for which the measured brightness has been erroneously increased through unresolved "optical" companions having been measured along with the variables. If the deviations are due to such a cause, the measured amplitude A' of a

variable so involved should be somewhat less than the true amplitude A; that is, the loss of amplitude through the effect of doubling is

$$\Delta A = (m_2 - m_1) - (m_2' - m_1')$$

where measured quantities are denoted by primes, and subscripts 1 and 2 refer to maximum and minimum light, respectively. If m_x is the magnitude of the superposed image, whether of one star or more, the true magnitude limits of the variable m_1, m_2, are found by familiar formulae or tables,[2] since the corresponding luminosities are $l_1 = l_1' - l_x$, $l_2 = l_2' - l_x$. The brighter the image x, the larger is ΔA.

The affected stars should, of course, occur proportionately more frequently in the most populous parts of the Cloud.

To test the suggested explanation of the large negative deviations, we have plotted amplitudes against residuals in figure 3. The data for variables in the nuclear region are plotted as circles, and in the border regions as crosses, when the residuals exceed ± 0.4. No correlation is obvious in this scatter diagram. The spread of amplitudes probably arises more from inherent causes than from superposed images.

The point is further examined in table 1, where mean amplitudes are given in specified intervals of deviation for the Cloud as a whole and for various regions of the Cloud. In what we call the nucleus, border and intermediate regions, the known variables have been studied with reasonable completeness;[3] the "other" variables are scattered throughout various other sections of the Cloud where the selection of variables for study has been somewhat superficial.

The similarity of the values of the mean amplitudes in the second column is striking. It means that in the Cloud as a whole the amplitudes are not related to the size and sign of the magnitude deviations from the period-luminosity curve. Equally striking is the similarity of mean amplitudes in nuclear, border and intermediate regions (last line of table 1). Location in the Cloud apparently does not systematically affect range of variation—either observationally, because of degree of crowding and nature of background, or physically, because of differences in gravitational field or other physical discriminants between nucleus and border of stellar systems.

Examining the various regions separately, however, we find that the results for the nucleus are strongly suggestive of a superposition effect. For all variables with $\Delta m < -0.09$, we find for the mean amplitudes

Nucleus	$0^m.824$,	36 stars
Borders	0.940	45 stars
Intermediate	0.930	63 stars
Other	0.990	103 stars
All	0.941	247 stars

TABLE 1

MEAN AMPLITUDES IN RELATION TO CLOUD STRUCTURE AND TO MAGNITUDE DISPERSION

RESIDUALS	ALL CLOUD	NUMBER	NUCLEUS	NUMBER	INTERMEDIATE	NUMBER	BORDER	NUMBER	OTHER	NUMBER
<-0.4	0.96	46	0.91	8	0.89	7	1.16	12	0.89	19
-0.39 to -0.22	0.95	99	0.76	9	0.91	25	0.85	20	1.05	45
-0.21 to -0.10	0.93	102	0.82	19	0.96	31	0.87	13	0.97	39
-0.09 to $+0.01$	0.96	97	0.92	15	0.95	28	0.99	21	0.98	33
$+0.02$ to $+0.14$	0.98	102	1.03	23	0.94	27	1.01	20	0.95	32
$+0.15$ to $+0.39$	0.96	86	0.98	13	0.93	23	0.94	26	1.01	24
$>+0.4$	0.96	30	1.05	10	0.86	6	0.83	9	1.15	5
All	0.956	562	0.93	97	0.93	147	0.95	121	0.99	197

and compare these values with the mean amplitude, $0^m.968$ (315), for all the other variables, $\Delta m > -0.9$, which presumably are not affected by doubling. Thus we have

$$\overline{\Delta A} = 0^m.968 - 0^m.824 = 0^m.144$$

for the nucleus (other regions appear unaffected). If this average difference in amplitude were assumed to be wholly produced by superposed images, the corresponding average deviation would need to be about $-0^m.25$, with the contributing star image, m_x, a magnitude and a half fainter than the variable at minimum. In the crowded star fields such superpositions seem entirely credible.

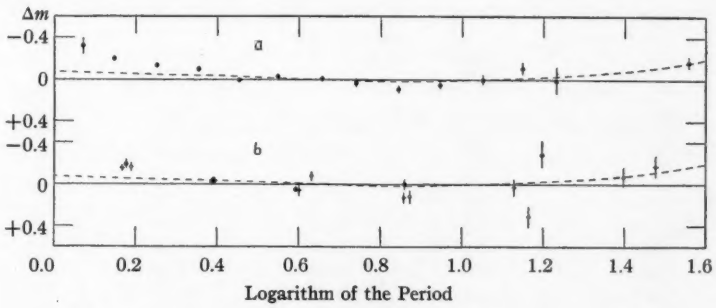

FIGURE 4

(*a*) Mean values for the material plotted in figure 2. Broken line represents adopted non-linear period-luminosity curve. Vertical lines through mean points indicate mean errors.

(*b*) Similar to (*a*) for selected regions of the Cloud. Circles refer to the nucleus, crosses to the border regions, dots to intermediate regions. The low point at log P = 1.164 includes some nuclear variables that are dimmed by obvious obscuration.

Since for stars of any given period there is probably a true dispersion in median luminosity, there may also be a dependence of amplitude on the median luminosity—a dependence which would work systematically to cancel or enhance any contribution to the scatter from superposed images. But unless such a dependence is operating in the sense that the more luminous variables (for a given period) have on the average greater amplitudes, and of just the amount necessary to cancel most of the effect from superposed images, we must conclude from the data of table 1 that except in the nuclear regions the contribution of doubling to the general dispersion in median magnitude is not important. This conclusion is supported by the appearance of the star fields, since the chances of trouble with superposed images, except for the fainter variables in the denser star fields of the

Cloud, seem to be slight for plates made with the Bruce refractor or the 60-inch reflector.

5. Test of the Period-Luminosity Curve.—In figure 4 (*a*) mean deviations are plotted against logarithms of the periods, for intervals of 0.1 in the logarithm, and are given also in table 2, along with the mean error of the mean deviation and the corresponding mean logarithm of the period. The mean errors are indicated in the figure by vertical lines through the points. The adopted non-linear period-luminosity curve has also been drawn in the figure. Although throughout its course it does not appear to be the best possible fit for this larger amount of material, it so nearly represents the most heavily weighted portion of the plot that a further revision at this time seems unnecessary and inadvisable. The lack of fit for the shortest periods and our reason for ignoring it have been mentioned above.

TABLE 2

RELATION OF MAGNITUDE DISPERSION TO LENGTH OF PERIOD

INTERVAL OF LOG OF PERIOD	MEAN LOG OF PERIOD	MEAN RESIDUAL	MEAN ERROR	NUMBER
0.0–0.1	0.071	$-0^m.32$	±0.08	5
0.1–0.2	0.148	−0.20	±0.04	37
0.2–0.3	0.252	−0.14	±0.02	67
0.3–0.4	0.355	−0.10	±0.03	59
0.4–0.5	0.454	0.00	±0.02	90
0.5–0.6	0.549	−0.03	±0.03	69
0.6–0.7	0.656	−0.01	±0.03	53
0.7–0.8	0.740	+0.03	±0.05	34
0.8–0.9	0.844	+0.09	±0.05	33
0.9–1.0	0.947	+0.05	±0.05	24
1.0–1.1	1.051	0.00	±0.06	21
1.1–1.2	1.149	−0.11	±0.07	23
1.2–1.3	1.233	+0.01	±0.14	16
> 1.3	1.559	−0.17	±0.06	33

6. Space Absorption Within the Small Cloud.—In table 3 we have tabulated mean magnitude deviations for intervals of 0.25 in the logarithm of the period for the nuclear, intermediate and border regions, separately. The numerous variables that do not lie in any one of these regions, where the survey has approached completeness, have not been included in this tabulation or in the corresponding figure 4 (*b*). Only 365 variables, rather than the 564 of figures 1, 2 and 4 (*a*), are therefore involved. Because of the small numbers of stars represented, the mean errors of some of the means are rather large. The adopted luminosity curve has been included, as in figure 4 (*a*).

If there were appreciable space absorption in the nuclear part of the Cloud, the apparent median magnitudes for stars in that region should be fainter than normal, and the plot in figure 4 (*b*) should reveal this system-

atic effect. But to the accuracy permitted by the mean errors of the plotted points, we can say that space absorption in the Small Magellanic Cloud is inappreciable. This conclusion does not mean that the Small

TABLE 3

PERIOD, MAGNITUDE DISPERSION AND CLOUD STRUCTURE

INTERVAL OF LOG OF PERIOD	MEAN LOG OF PERIOD	MEAN RESIDUAL	MEAN ERROR	NUMBER
		Nucleus		
0.00–0.25	0.190	−0.16	±0.05	12
0.25–0.50	0.386	−0.02	±0.03	23
0.50–0.75	0.631	−0.08	±0.06	22
0.75–1.00	0.874	+0.12	±0.07	21
1.00–1.25	1.164	+0.31	±0.11	12
> 1.25	1.398	−0.07	±0.10	7
		Intermediate		
0.00–0.25	0.178	−0.19	±0.05	21
0.25–0.50	0.391	−0.04	±0.03	65
0.50–0.75	0.593	+0.05	±0.03	38
0.75–1.00	0.860	0.00	±0.06	16
> 1.00	1.198	−0.28	±0.14	7
		Border		
0.00–0.25	0.168	−0.15	±0.03	21
0.25–0.50	0.394	−0.03	±0.04	39
0.50–0.75	0.601	+0.06	±0.07	21
0.75–1.00	0.858	+0.13	±0.06	15
1.00–1.25	1.129	+0.03	±0.09	16
> 1.25	1.476	−0.17	±0.11	10

Cloud is wholly free of absorbing material; a few bright and dark patches have been observed. Nor does it refer to the space absorption between the observer and the Small Cloud. That absorption, which is believed to amount to three-tenths of a magnitude in photographic light, has been indirectly detected through the scarcity on Harvard long-exposure plates of remote external galaxies in the direction of the Cloud.

[1] Eighth paper of this series, these PROCEEDINGS, **26**, 543–544 (1940).

[2] For example, *Harv. Ann.*, **33**, 287 (1900).

[3] See fifth paper of this series, these PROCEEDINGS, **26**, 110 (1940).

GALACTIC AND EXTRAGALACTIC STUDIES, XV. ON THE DISTRIBUTION OF PERIODS FOR 343 CEPHEIDS IN THE SMALL MAGELLANIC CLOUD

BY HARLOW SHAPLEY AND VIRGINIA MCKIBBEN

HARVARD COLLEGE OBSERVATORY

Communicated April 13, 1942

In the present communication we treat further the distribution of periods of classical Cepheids, a subject that has already been considered in the fifth paper of this series[1] for the Magellanic Clouds and in the eleventh[2] for the nucleus of our own Galaxy.

In both of the earlier discussions it was found that the conventional period-frequency curve, which is based on the galactic Cepheids that are chiefly in the solar neighborhood, is apparently of only local applicability. Not only does it fail to represent the distribution of periods found in the Magellanic Clouds, but it fails also at the galactic nucleus and elsewhere. Moreover, because of the dependence of period on absolute magnitude, and the consequent favoring of the longer periods in any survey based on apparent magnitude, the solar neighborhood cannot be impartially explored; and therefore the conventional curve of period frequency is not accurately representative even of the space around the sun. An appropriate correction for this absolute magnitude selection changes the local curve toward a resemblance to the frequency curve for the Small Cloud, but only partly cancels the observed differences.

We also give in the present paper a preliminary determination of the distance of the "wing" of the Small Magellanic Cloud, a comment on the classical Cepheids in the globular clusters, and a sample of the light curves of the "16-day" Cepheids.

1. The Distribution of Periods.—Of the 564 classical Cepheids that have now been studied in the Small Cloud, and for which periods and magnitudes have been discussed in the preceding paper of the series, 343 lie within the nuclear, intermediate and border regions, where the search for and measurement of the variables have been reasonably complete. The other 221 variables are outside these regions, in parts of the Cloud where the survey has been incomplete and where a preferential selection of the brighter stars and the longer periods has no doubt operated. These outside variables are obviously not suitable for a study of the period frequency; in fact, they give, in their incomplete representation, a distribution curve that rather closely resembles the curve for the 288 galactic Cepheids.[1]

The positions in the Cloud of the chosen nuclear, intermediate and border regions are the same as specified in the earlier paper, but all have been enlarged. It is believed that by using only the results in these

thoroughly worked regions we are getting the most nearly correct picture of the distribution of the periods of classical Cepheids in this particular galaxy; and perhaps it is the correct picture of the normal or average behavior of Cepheids in our Galaxy, when it is considered as a whole.

TABLE 1

DISTRIBUTION OF PERIODS OF CEPHEIDS IN SPECIAL SECTIONS OF THE SMALL MAGELLANIC CLOUD

REGION	NUMBER	PERCENTAGE OF VARIABLES WITH PERIOD				
		$<2^d$ %	2^d TO 10^d %	$>10^d$ %	14^d TO 17^d %	$>20^d$ %
Nucleus	74	10.8	63.5	25.7	10.8	8.1
Core	24	8.3	50.0	41.7	16.7	16.7
Intermediate	147	21.8	73.5	4.8	0.7	0.7
Borders	122	23.8	54.9	21.3	5.7	7.4
All	564	19.3	64.2	16.5	3.7	5.9

Peculiar distributions occur in special localities, such as the galactic nucleus and globular clusters, or at great distances above and below the galactic plane, where with rare exceptions Cepheids appear only in the form of cluster-type variables.

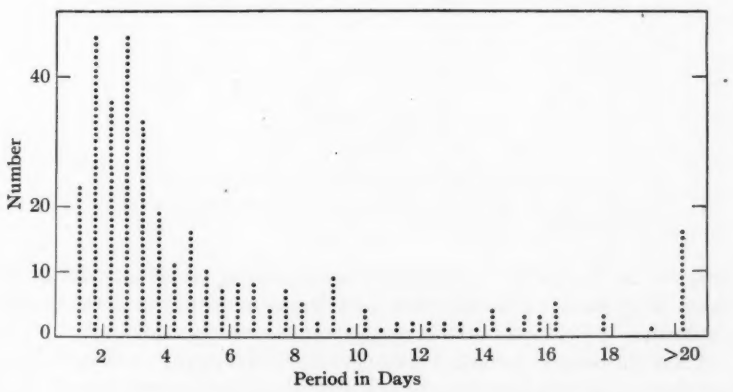

FIGURE 1

Period distribution of Cepheids in the Small Magellanic Cloud.

The material from the three regions is brought together in figure 1, which shows the number of variables for each half-day interval in the period. Nearly seventy per cent of the variables have periods less than 3.5 days, in striking contrast to only twelve per cent in our galactic system. The maximum of the frequency curve in the Small Cloud is close to two days; in our own system, 4.5 days. There are no certain secondary

maxima in the curves for the Cloud, although we note a considerable number of periods between fourteen and eighteen days. Periods of this length are also rather common among galactic Cepheids, and the unusual group in the direction of the galactic center is mentioned below.

In table 1 some data on the frequency of periods in the different regions of the Cloud are assembled. In making this tabulation twenty-three periods based on the study of plates made with the 60-inch reflector have been withdrawn from the material available for the nucleus because similar work with the reflector has not yet been completed for the borders and the intermediate regions. The larger scale of the reflector plates and their fainter magnitude limits permit the discovery of variables that have been overlooked on the Bruce plates.

Although the reflector plates reach well below the limit where cluster-type variables should appear, so far not one has been definitely found and

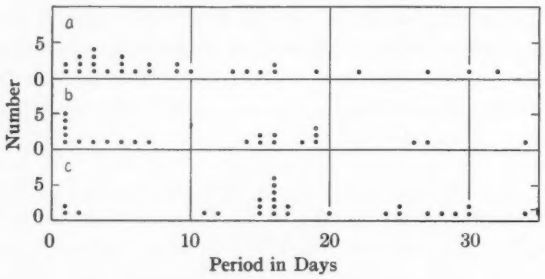

FIGURE 2

Period distribution of Cepheids in: (*a*) Core of Small Magellanic Cloud; (*b*) Globular clusters; (*c*) Galactic nucleus (longest period is 69.05 days).

assigned to the Cloud. It appears that if cluster variables exist in the Small Magellanic Cloud they are very scarce in comparison to classical Cepheids.

A few cluster-type variables are found on the photographs of the Cloud, but they are all brighter than the seventeenth magnitude and are not more abundant than would be expected for the foreground. Similarly a number of eclipsing stars, which probably are in the foreground rather than giant binaries in the Cloud, have been found and studied. It is difficult to be as certain concerning the non-membership of the several long-period variables. If they are in the Cloud they are of the supergiant type such as recorded for 47 Tucanae. Some of them certainly are foreground stars—typical long-period variables of our own galactic system.

2. Anomalous Distributions of Period.—The densest portion of the nucleus of the Small Magellanic Cloud has been called the "core." Its

variables are listed separately in table 1. The great contrast in the distribution of periods between this core and the border regions was also noted in the earlier discussion. The cylindrical core extends, of course, from the near border to the far border, and the true contrast between the actual center of the Cloud and the border is therefore unattainable; we cannot tell which of the core or nuclear variables are really near the center,

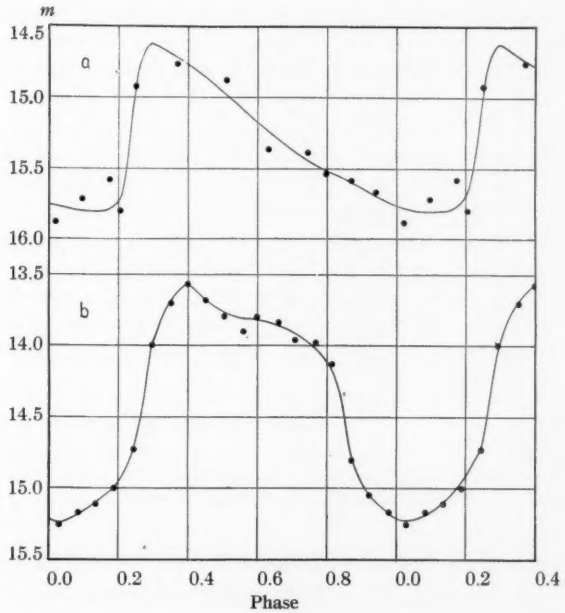

FIGURE 3

Light curves of two 16-day Cepheids: (*a*) HV 1333 and (*b*) HV 10495. Ordinates are photographic magnitudes; abscissae are fractions of the periods.

which far out in the cylinder. Consequently the contrast observed is a minimum.

The distribution for the core is shown graphically in figure 2, and in the same diagram is given the distribution for the twenty-three classical Cepheids in globular clusters,[3] and for the twenty-seven in the region of the galactic center. The data for the nucleus of the Galaxy are taken from the first table of the eleventh paper of this series, after removing five stars that are known from photometric data to be in the solar neighborhood and far this side of the center of the system. A comparison of figures 1 and 2

emphasizes the peculiar association of at least some of the long period Cepheids with the densest star fields; or perhaps the more striking point is the almost complete absence from these crowded regions of the average sort of Cepheid, with a period from two to seven days.

3. Light Curves for the Sixteen-Day Cepheids.—For convenience we call the Cepheids with periods between fourteen and seventeen days the "16-day" Cepheids. In the Small Magellanic Cloud these stars all appear to have typical Cepheid light curves, such as that suggested in figure 3 (*a*) for HV 1333. But about one-half of those in the globular clusters and at the galactic center have light curves much like that shown in figure 3 (*b*) for HV 10495. This uncommon light curve is like that of W Virginis,[4] a variable of period 17.27 days, which is, incidentally, in a very low density region far from the galactic plane. Much careful photometry and spectroscopy should be done on these 16-day Cepheids because of their possible significance in the study of the energy sources in supergiant stars.

4. The Distance of the Small Cloud's Wing.—A faint extension of the Small Magellanic Cloud in the general direction of the Large Cloud has been reported in earlier papers.[5] Whether this wing is actually a part of the Small Cloud cannot be told from general appearances, but it is now possible to apply the Cepheid test. The following four variables fall in the area of the wing:

NAME OF VARIABLE	PERIOD IN DAYS	MEDIAN MAGNITUDE	AMPLITUDE OF VARIATION
HV 2233	15.172	$14^m.4$	$1^m.25$
VMcK 14	16.2	15.35	0.9
EHB 37	21.4	14.5	1.4
HV 865	33.3	14.0	1.35:

The means of the median magnitudes and the logarithms of the period are 14.56 ± 0.30 and 1.31 ± 0.08, respectively, and the corresponding distance modulus is

$$m - M = 17.18 \pm 0.34 \text{ (m.e.)}$$

This value should be compared with the distance moduli 17.10 and 17.35 for the Large Cloud and Small Cloud, respectively. To the accuracy permitted by the scanty data, we can conclude that the wing is probably a part of the Small Magellanic Cloud. Available photographs indicate that it does not have a dense background of faint stars, but there are a number of faint star clusters within its bounds, and some patches of diffuse nebulosity.

[1] These PROCEEDINGS, **26**, 105–115 (1940); *Harvard Reprint* 192.
[2] Op. cit., **26**, 681–688 (1940); *Harvard Reprint* 214.
[3] Miss Sawyer, *Publ. David Dunlap Obs.*, **1**, No. 4 (1939).
[4] Gaposchkin, *Harv. Bull.*, 906, 8 (1937).
[5] *Harv. Bull.*, 914, 8 (1940); *The Telescope*, **8**, 15 (1941).

CLOSE LINKAGE BETWEEN MUTATIONS WITH SIMILAR EFFECTS

BY L. C. DUNN* AND ERNST CASPARI†

COLUMBIA UNIVERSITY AND LAFAYETTE COLLEGE

Communicated April 2, 1942

About a dozen different mutations which affect the tail and other parts of the axial skeleton of the house mouse have been reported by various observers. Five of these appear to form a group of which the members are related to each other, first, by producing similar effects on the tail and, second, by showing little or no crossing-over with each other. The genetic relations among these five have now been tested with the result that three of them behave as though unilocal, while two others show less than 3% of recombination with the first locus or with each other. It appears thus that three adjacent loci have each produced mutations of similar kinds. This clustering of five similar mutations indicates a relationship between the location and the morphological effect of mutations which deserves careful study. A summary of the crossing-over data involving the five mutations is given herewith to be followed by a fuller discussion elsewhere.

The mutations involved are listed below, with references to the original or most complete description of their effects.

T, dominant Brachyury; tail short or absent; varies in different stocks; occasionally has fused or reduced vertebrae (Dobrovolskaia-Zawadskaia, 1934); *TT* is lethal (Chesley, 1935).

Fu (= T^f) dominant Fused; fused and reduced vertebrae; moderately shortened; *Fu Fu* is viable (Reed, 1937).

Ki, dominant Kink; effect like *Fu*; *Ki Ki* is lethal (Caspari and David, 1940).

t^0, recessive lethal; Tt^0 is tailless (Chesley and Dunn, 1936).

t^1, recessive lethal; Tt^1 is tailless (Dunn and Gluecksohn-Schoenheimer, 1938).

(t^2, recessive lethal like t^1; Tt^2 is tailless, t^1t^2 is lethal (Dunn, 1939) discarded; and not used in further work.)

All combinations of the five mutations have now been observed and a study of the phenotypic effects permits the following statements. *T*, *Fu* and *Ki* are clearly dominant in all combinations and have effects on the tail which show overlapping variability so that the three mutants cannot be reliably distinguished when heterozygous. The three dominants differ in their reactions with t^0 and t^1, since Tt^0 and Tt^1 are tailless while *Fu* t^0, *Fu* t^1, *Ki* t^0, and *Ki* t^1 all have deformed tails of the same sort, that is, like Brachy or Fused. On this basis, we might place *Fu* and *Ki* in one group, and *T*, t^0 and t^1 in another. Phenotypically there is further justification for such a grouping since animals with either *Fu* or *Ki* often show waltzing or choreic

behavior whereas these characters are not parts of the phenotypic effect of T, t^0 or t^1. T, *Fu* and *Ki* produce similar phenotypes in combination with each other. Combinations of T with *Fu*, T with *Ki* and *Fu* with *Ki* may have somewhat more extreme effects on the tail than the separate mutations, but the material is too variable to permit exact statements. On the basis of mutual interaction effects we might then recognize two groups: (1) *Fu* and *Ki*, (2) T, t^1 and t^0.

In genetic behavior no such separation into groups has heretofore been possible. T, t^0 and t^1 have shown no crossing-over among themselves in over 6000 observations (Dunn unpublished). Reed (1937) did not observe any certain crossovers between *Fu* and T among 240 offspring from matings of $\frac{Fu\ +}{+\ T}$ by normal. The tests were complicated by the fact that *Fu* frequently failed to manifest itself, necessitating progeny tests of all suspected crossovers.

We have carried out new experiments designed to detect crossing-over between *Fu* and T (data chiefly from Dunn), *Ki* and T, *Ki* and *Fu* (data chiefly from Caspari); between *Ki*, t^0 and t^1 and *Fu*, t^0 and t^1 (Dunn).

The results for *Fu* and T are shown in table 1.

TABLE 1

PARENTS	OFFSPRING: ABNORMAL TAIL T OR *Fu*	TAILLESS	NORMAL TAIL	NORMALS TESTED *Fu* (OVERLAPS)	NORMALS TESTED ++ (RECOMBINATIONS)
$++ X \frac{Fu\ +}{+\ T}$	356	4	13	6	4
$\frac{Fu\ +}{+\ T} X ++$	30	..	2	..	..
Total	386	4	15	6	4

The normal-tailed offspring arise either from recombination between *Fu* and T or from failure of one of these dominants to express itself in the offspring. Ten of the normal offspring have been sufficiently tested by crossing with normals. Six of them have produced both Fused and normal progeny showing them to be $Fu\ +$; four of them, however, have produced, respectively, 37, 37, 33, 26 normal-tailed offspring, proving them to be $++$. Out of a total of 400 offspring (386 *Fu* or T, 4 tailless and 10 tested normals) a minimum of four recombinations of the $++$ type occurred.‡ If the test cross was $\frac{Fu\ +}{+\ T} X++$, a contrary type of crossover, $\frac{Fu\ T}{++}$, is to be expected in equal numbers. These have not yet been detected, although it is likely that they are to be sought among the tailless animals which are weak and often sterile. The amount of recombination detected therefore repre-

sents half of the expected amount. The amount of recombination between *Fu* and *T* may thus be estimated roughly at about 2.0% (8/400).

From *inter se* matings of animals known to be $\frac{Fu\ +}{+\ t^0}$, 123 Fused and 10 normal offspring have been obtained. Tests of the normals have not been completed. Three of them are known to be $\frac{Fu\ +}{+\ t^0}$ (overlaps).

We have carried out similar tests with Kink and Brachy. At Columbia, four males and two females heterozygous for these two mutations $\left(\frac{Ki\ +}{+\ T}\right)$ were crossed with normal animals (++) and have given 408 offspring of which 406 were either Kink or Brachy, one male had a normal tail and one was tailless. The normal-tailed male when tested by normal-tailed females gave 39 offspring, all normal and was thus a ++ recombination (or a mutant).

The tailless male when tested by normal-tailed females gave *18 tailless*, 21 normal and 10 Kink or Brachy. This suggested that this exceptional male was $\frac{Ki\ T}{+\ +}$, that is, a recombination in which *Ki* and *T* were in the same chromosome. This was proved by mating him to Brachy females (*T* +) and dissecting the mothers on the eleventh day of pregnancy. Typical *T T* homozygotes (cf. Chesley, 1935) were found. We are grateful to Dr. Gluecksohn-Schoenheimer for making these observations. Some of the abnormal-tailed offspring from matings of this male by normal females showed the waltzing behavior and deafness which occurs with *Ki* but never with *T* alone. These facts show him to contain *T* and most probably *Ki* in the arrangement $\frac{T\ Ki}{+\ +}$. The fact that he was tailless (the only tailless animal out of 785 from similar ancestry) suggests that *Ki* and *T* may produce a greater interaction when both are present on the same chromosome.

The $\frac{Ki\ +}{+\ T}$ males tested by $\frac{+\ t^0}{+\ t^1}$ females have given 192 with abnormal tails (98 *Ki*, 94 tailless Tt^0 or Tt^1) and 8 with normal tails. Three of the latter have proved to be $+t^1$ or $+t^0$ and are thus due to recombination, while three proved to carry *Ki*.

At Lafayette the result of crossing heterozygous females by normal males was 124 Kink or Brachy and 8 normal. The reciprocal cross yielded 245 Kink or Brachy and 8 normal. Of the 8 normals tested, 6 proved to be ++. The total number of ++ recombinants is thus 10. Assuming these to represent half the total, the estimated frequency of recombination between loci *T* and *Ki* is about 2.0% (24/986).

The data are summarized in table 2.

TABLE 2

PARENTS	OFFSPRING: ABNORMAL TAIL (T OR Ki)	NORMAL TAIL	NORMALS TESTED Ki (OVERLAPS)	NORMALS TESTED $++$ (RECOMBINATIONS)	TAILLESS TESTED $\frac{Ki\ T}{+\ +}$
$++ X \frac{Ki +}{+ T}$ (C)	381	1	..	1*	1
$\frac{Ki +}{+ T} X ++$ (C)	25	..	..	..	..
$++ X \frac{Ki +}{+ T}$ (L)	245	8	2	3	..
$\frac{Ki +}{+ T} X ++$ (L)	124	8	2	5	..
$\frac{+ t^0}{+ t^1} X \frac{Ki +}{+ T}$ (C)	192	9	3	3†	..
Totals	967	26	7	12	1

(C). At Columbia.

(L). At Lafayette.

* One male tested by $++$ gave 39 normal progeny.

† Each animal tested by $++$ gave 25 or more normal progeny. One male tested by $++$ gave 30 normal and by $T+$ gave 12 normal, 1 Brachy, 8 tailless, hence is $+ t^1$ or $+ t^0$. One male tested by $++$ gave 43 normal and by $T +$ gave 12 normal, 9 Brachy, and 5 tailless, hence is $+ t^0$ or $+ t^1$; one male tested by normal, gave 25 normal and by $T+$ gave 8 normal, 6 Brachy, 9 tailless.

From matings of $\frac{Ki\ +}{+\ t^1} X \frac{Ki\ +}{+\ t^1}$ 93 offspring have been recorded of which 91 were Kink and 2 were normal. The latter have not yet been tested. Because of the close linkage between the two lethals *Ki* and t^1, the stock containing both may be maintained as a balanced lethal stock by eliminating the rare normals which arise by recombination.

It has also been possible to test whether *Ki* and *Fu* are mutations at the same or different loci. Animals of genotype $\frac{Fu +}{+ Ki}$ (having the appearance of Fused or Kink) crossed with normal-tailed testers gave 172 *Fu* or *Ki* and 21 normals. Of 11 of the latter, one when tested gave only normal-tailed offspring (61 normal test offspring and 1 with slightly abnormal tail). Another one may be a recombinant, too, having had 29 normal progeny only. The other 9 are overlaps.

These results appear to us to establish the fact that *T*, *Ki* and *Fu* are mutations at three separate loci. The relative recombination frequencies and the order of the loci cannot be established from the present limited data in which the amount of recombination represents a minimum estimate, since some exceptional animals die before sufficient offspring can be ob-

tained to classify them definitely as recombinants or overlaps. It is nevertheless evident that *T*, *Ki* and *Fu* are closely linked, probably all within less than three crossover units. The mutations at these three loci are thus not only phenotypically similar but seem to be near neighbors in the chromosome. We are led, therefore, to inquire whether their similarities and their proximity in the chromosome may be related.

It is conceivable, of course, that the two circumstances have no causal relation at all; and that by chance the mutations occurring at three adjacent loci happen to have similar effects. The likelihood of such a coincidence can be estimated by assuming the ten loci in the mouse known to have produced mutations of a similar type§ to be distributed at random among the 20 chromosomes. The chance that any chromosome should have one of these loci is thus 10/20 or $^1/_2$; the chance that one chromosome should have 3 such loci is $(^1/_2)^3$. The genetic lengths, in crossing-over units, of the chromosomes of the mouse have not been determined. However, the average number of chiasmata per chromosome may be estimated as just over two (Crew and Koller, 1932) and this would lead to an estimated average map length of about 100. Assuming, therefore, that a chromosome consists of 25 crossover segments of about four crossover units each, and that crossing-over in this chromosome is normal, the chance that three loci chosen at random should fall within one segment would be $(^1/_{25})^3$. Consequently, the likelihood that three loci should have the distribution actually found in our sample would be $(^1/_2)^3 \cdot (^1/_{25})^3$ or 0.000008. This is rather remote; and it would seem more likely that the similar effects and near location are related to each other.

The simplest assumption concerning this relationship would be that *T*, *Fu* and *Ki* represent mutations in duplicate loci lying next to each other in the chromosome in the manner of the "repeats" discovered in Drosophila by Bridges (1935). Lewis (1942) has recently shown that Star and asteroid, two phenotypically similar and very closely linked mutations in *Drosophila melanogaster* are located in one double band of the salivary chromosome, the doubling having presumably occurred through repetition of one band. Although such a repeat hypothesis would reconcile the phenotypic similarity and the close linkage existing among the three dominants, it fails to account for the recessives at the *T* locus (t^0, t^1) and the fact that $t^0\,t^0$, $t^1\,t^1$, *T T* and *Ki Ki* are lethal; that is, these recessive effects are not "covered" as they should be if normal duplicate loci are present. Although the duplication hypothesis might be rescued by assuming that the duplicated normal loci have lost, by virtue of their position, that portion of their normal effect which "covers" the lethal effect, we have no independent evidence that this is so. It is possible, however, to test one consequence of the duplication hypothesis, namely, that mutations at repeated or duplicated loci should retain some basic similarity in developmental effect like

that which marks the mutations at the same locus. To this end, the effects of *Fu* and *Ki* on embryological processes are being compared with those already known to be associated with *T*, t^0 and t^1 (see Dunn, 1941, for review).

It is, of course, possible to conceive of a developmental unit in the chromosome having greater extension than one crossing-over unit and to regard the manifestations of these several mutations as effects of such a larger unit. However, until it is known whether the developmental effects are unitary or diverse such ideas will lack support. It will also be necessary to test whether crossing-over is normal in the chromosome involved, since it is possible that the small amount of recombination between loci *T*, *Fu* and *Ki* is due to some more general suppression of crossing-over. If the developmental processes chiefly affected turn out to be not as similar as the adult phenotypes seem to indicate or if crossing-over is not normal, then chance association of three similar mutations would again have to be considered as a possible explanation.

* It is a pleasure to acknowledge the help of Dr. Salome Gluecksohn-Schoenheimer and the financial assistance of the fund for research of Columbia University.

† Aided by a grant of the John and Mary R. Markle Foundation.

‡ Since this shows Brachy and Fused to be at separate loci, we shall adopt the symbol *Fu* replacing T^F for the Fused locus, retaining *T* for Brachy.

§ The ten loci are: *T*, t^0, t^1 (Dunn, 1941); *Fu* (this paper); *Ki* (Caspari and David, 1940); *Sd* (Dunn, Gluecksohn-Schoenheimer and Bryson, 1940); *st* (Dunn, 1934); *sb* (stub) (Dunn and Gluecksohn-Schoenheimer (in press)); *tw* (twist) (Dunn and Gluecksohn-Schoenheimer, unpublished); *sc* (screw) (Laanes and MacDowell, 1942); *fl* (flex) (Hunt, *et al.*, 1933); *pt* (pigtail) (Crew, 1941).

Bridges, C. B., *Jour. Heredity*, **26**, 60–64 (1935).

Caspari, E., and David, P. R., *Ibid.*, **31**, 427–431 (1940).

Chesley, P., *Jour. Exp. Zool.*, **70**, 429–459 (1935).

Chesley, P., and Dunn, L. C., *Genetics*, **21**, 525–536 (1936).

Crew, F. A. E., and Auerbach, C., *Jour. Genetics*, **41**, 267–274 (1941).

Crew, F. A. E., and Koller, P. Ch., *Ibid.*, **26**, 359–383 (1932).

Dobrovolskaia-Zavadskaia, Kobozieff, N., and Veretennikoff, S., *Archives de Zoologie Experimentale*, 1934, pp. 250–358.

Dunn, L. C., these PROCEEDINGS, **20**, 230–232 (1934).

Dunn, L. C., *Genetics*, **24**, 728–731 (1939).

Dunn, L. C., *Third Growth Symposium*, 1941, pp. 147–161.

Dunn, L. C., and Gluecksohn-Schoenheimer, S., *Genetics*, **23**, 146–147 (1938).

Dunn, L. C., Gluecksohn-Schoenheimer, S., and Bryson, V., *Jour. Heredity*, **31**, 343–348 (1940).

Hunt, H. R., Mixter, R., and Permar, D., *Genetics*, **18**, 335–366 (1933).

Laanes, T., and MacDowell, E. C., *Ibid.*, **27**, 151–152 (1942).

Lewis, E. B., *Ibid.*, **27**, 153–154 (1942).

Reed, S. C., *Ibid.*, **22**, 1–13 (1937).

THE EFFECT OF pH ON INACTIVATION OF TOBACCO MOSAIC VIRUS BY X-RAYS*

BY A. MARSHAK† AND WILLIAM N. TAKAHASHI

RADIATION LABORATORY AND THE DIVISION OF PLANT PATHOLOGY, UNIVERSITY OF CALIFORNIA

Communicated April 13, 1942

Previous experiments have shown that the sensitivity of chromosomes to x-rays can be reduced by treatment with ammonium hydroxide and that the protective action of the ammonium hydroxide increases with concentration.[6] This effect was attributed to the removal of positive charges on the chromosome surfaces. The experiments with tobacco mosaic virus reported here show that on the acid side of the isoelectric point where the molecules carry a net positive charge they are more sensitive to x-rays.

Experimental.—The purified virus, very kindly given to us by Dr. W. M. Stanley, was isolated by ultracentrifugation. One-tenth of a cc. of a distilled water solution containing 2.1 mg. of the virus was suspended in nine-tenths of MacIllvaine buffer at the proper pH and irradiated in paraffin-lined celluloid capsules. One sample of virus in buffer at pH 7.0 (reference point) was always irradiated at the same time as samples at other pH values. After irradiation the volume of solution was made up to 100 cc. and adjusted to pH 7.0 with the appropriate buffer solution. Controls received the same treatment except that irradiation was omitted. Experiments were limited to pH values from 2.2–7.0 since investigations of other workers[9] have shown that the virus is inactivated beyond that range.

The relative concentration of active virus was determined by the usual biological method of inoculating opposite halves of *Nicotiana glutinosa* leaves[4, 7] with the irradiated virus and unirradiated controls. For each sample 20 half-leaves were used.

In a preliminary experiment portions of a single sample of virus in distilled water were given different doses of x-rays and tested on half-leaves after diluting 1/1000. The curve of the logarithm of the number of lesions produced by irradiated virus as plotted against per cent of the unirradiated control is given in figure 1. The results fit the equation $Y = e^{-kx}$, where Y is the per cent unaltered virus, k a constant and x the dose in Roentgens. k has the value of 8.12×10^{-6} and the six points obtained fit the curve to within 1%. The sample of virus irradiated two weeks later showed inactivation within 2% of the values previously obtained. However, another sample of virus prepared in the same manner showed much greater sensitivity to x-rays. Curve *a* of figure 1 gives the results obtained with the first sample and curve *b* those of the second. The slopes of the two curves and therefore the x-ray sensitivity of the two samples of virus

differ by a factor of approximately 9. It was found necessary to use different samples of virus in the course of the first pH experiment. This, as will be seen later, did not influence the results obtained.

Results.—The counts of local lesions on pairs of half-leaves are listed in table 1.

TABLE 1

N. glutinosa LESION COUNTS FOLLOWING IRRADIATION AT DIFFERENT pH VALUES

	(1)		(2)		(3)		(4)	
pH	C_7	I_7	C_x	I_x	C_7	C_x	I_7	I_x
2.2	2015	1334	2025	1023	2430	2265	1002	613
3.0	2066	1344	2255	1387	2602	2457	1097	922
4.0	1213	867	1397	1017	(1)1154	1010	744	702
					(2) 433	442		
5.0	1854	1361	1537	1048	1486	1312	1308	1229
6.0	1854	1361	1707	1344	1611	1456	1445	1442

Column 1 gives the number of lesions produced by unirradiated virus at pH 7.0 (C_7) and irradiated virus at pH 7.0 (I_7), each being on halves of the same leaf of *N. glutinosa*. Similarly the spots produced on one set of half-leaves by unirradiated virus at different pH values is given by C_x (column 2) and the spots produced on corresponding half-leaves by virus irradiated at different pH concentrations by I_x. The controls in column 3 are compared with the irradiated samples in column 4.

The fraction of non-inactivated virus at pH 7.0 for any one run is given by I_7/C_7 and similarly at other pH concentrations by I_x/C_x. Letting $I_7/C_7 = S_7$ and $I_x/C_x = S_x$, then S_x/S_7 will give the survivors at pH_x as a fraction of the survivors at pH_7, and $1-S_x/S_7$ will give the efficiency of the irradiation in inactivating the virus at pH_x. The results of these computations are given in column *a* table 2. The efficiency of the x-rays at different pH concentrations may also be calculated from the data of (3) and (4) of table 1. The virus inactivation by pH alone will be given by $1 - (C_x/C_7)$ and this subtracted from $1 - (I_x/I_7)$ will give the x-ray efficiency of pH_x.

TABLE 2

THE RELATIVE EFFICIENCY OF X-RAYS AT DIFFERENT pH VALUES

	a	*b*
pH	$\left(1-\frac{S_x}{S_7}\right)$ 100	$\left[\left(1-\frac{I_x}{I_7}\right)-\left(1-\frac{C_x}{C_7}\right)\right]$ 100
2.2	+23	+32
3.0	+10	+10
4.0	− 2	− 7, +6
5.0	+ 7	− 6

Table 2 gives a summary of the x-ray efficiency in per cent of the inactivation at pH 7.0. At pH 4.0 a second set of inoculations was made with C_x and C_7 with the result shown in table 1. It is clear from this and

from the differences in the values obtained by the two methods described that errors as large as 13% may be expected.

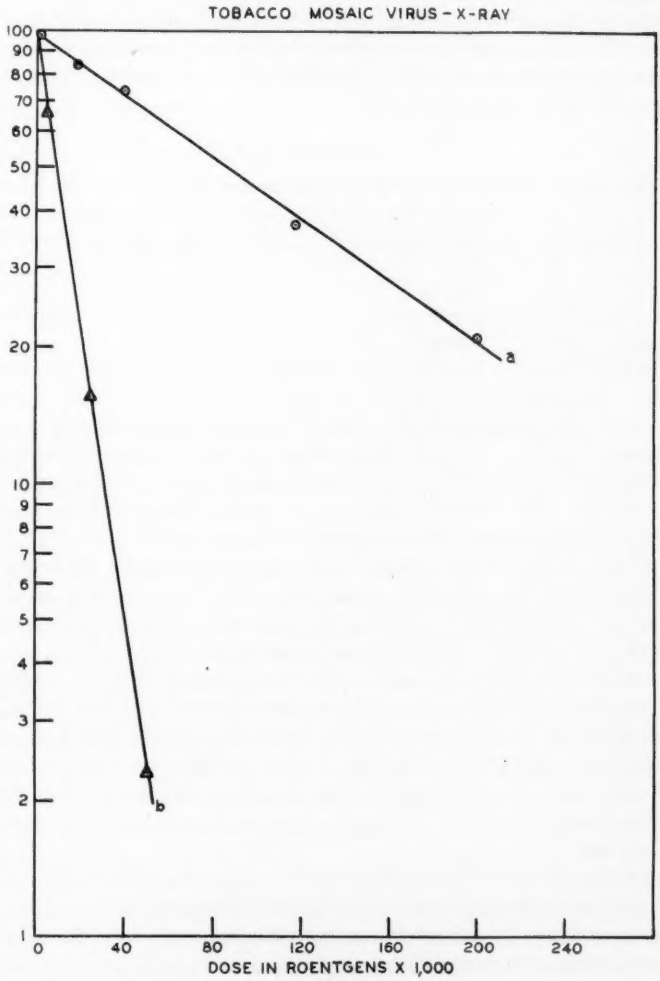

FIGURE 1

From examination of table 2 it appears that the efficiency of x-rays between pH 4.0 and 6.0 is equal to or less than the efficiency at pH 7.0. The differences are not large enough to be significant. However, at pH 2.2–3.0 the efficiency is significantly increased. To determine whether this might

have been an accidental result two more experiments were performed. In both of these the same sample of virus was used in determining inactivation at all pH values used. The results shown in table 3 are essentially similar to those of the previous experiment. At pH 2.2 x-rays are 25–30% more efficient in inactivating the virus than at pH 7.0, and at pH 3.4 the x-ray efficiency is greater by 10–15%. Between pH 3.4 and pH 7.0 there are no significant differences in the results.

TABLE 3

N. glutinosa LESION COUNTS AND X-RAY EFFICIENCY AT DIFFERENT pH VALUES

		$C_{2.2}$	$I_{2.2}$	$C_{3.4}$	$I_{3.4}$	$C_{5.6}$	$I_{5.6}$	$C_{7.0}$	$I_{7.0}$
Number	1	608	276	515	280	298	207	425	276
	2	220	110	266	156	322	207	284	199
$\frac{I}{C}$	1	0.454		0.544		0.695		0.650	
	2	0.50		0.586		0.643		0.678	
$\left(1-\frac{S_x}{S_7}\right)$ 100	1	+30		+16		−7		0	
	2	+26		+13		+5		0	

Gowen[3] studied the effect of x-rays on tobacco mosaic virus and found exponential survival curves. From these curves he calculated the size of the virus to be 7.5×10^{-18} cm.^{3}. He does not report differences in survival curves with different samples of virus. The sensitive volumes calculated from curves *a* and *b* of our experiments are 4.6×10^{-18} cm.^{3} and 4.2×10^{-17} cm.^{3}, respectively. Lea and Smith[5] irradiated dried and aqueous suspensions of tobacco necrosis virus with x-rays and found the results of both methods to fall on the same exponential survival curves. They mention that in different experiments different rates of inactivation were observed and the comparison mentioned above is made only with what they consider their best results. They found that tobacco mosaic virus in different states of aggregation gave essentially similar survival curves when treated with ultra-violet light. However, the experimental error of their observations was sufficiently large to make it impossible to determine whether their data indicated a response by elementary particles or aggregates as high as 4.

In the present experiments the shape of the survival curve is determined with an error no greater than 1–2% and can be explained only in terms of a response to x-rays by a single functional virus unit. Pirie[1] and Frampton[2] have presented evidence indicating that the virus particles may exist as aggregates. Filtration studies of the tobacco necrosis virus by Smith and MacClement[6] indicate the presence of aggregates in extracts containing this virus. The form of the survival curve we have obtained requires that at least one functional unit of the virus be inactivated for every ion pair or cluster produced in the virus elementary particle or aggregate. If aggregates exist, the following possible interpretations may be given to the data:

(1) There is only one functional unit in the aggregate; the rest is inert material.

(2) Aggregates are not permanent but are continually being broken down and built up from elementary units. As pointed out by Lea and Smith in this case the inactivation of particles in aggregates will proceed at the same rate as the inactivation of elementary particles and simple exponential curves will be obtained.

In both (1) and (2) we may assume that:

(*a*) The energy released by an ion pair or cluster within one elementary unit will inactivate only that unit.

(*b*) The energy released by an ion pair or cluster will inactivate more than one unit.

If after irradiation the virus is diluted before inoculation the number of aggregates will be decreased. Assuming condition (1), the apparent sensitivity of the virus will be independent of the state of aggregation and the spread of energy through the aggregate could not be detected. Under condition (2*a*) the sensitivity of the virus will be independent of the state of aggregation, while under condition (2*b*) the greater the aggregation at the time of the irradiation the greater the apparent sensitivity. The results obtained point to the latter hypothesis.

If it is postulated that the greater sensitivity to x-rays at pH 2.2 to 3.4 is due to a greater number of aggregates than at pH 3.4–7.0 the results cannot be adequately explained. One would expect a maximum sensitivity at pH 3.4, the isoelectric point, whereas the maximum observed is at 2.2. One would also expect that the samples of virus containing different amounts of aggregates would have different relative responses to x-rays at the various pH values. However, the same response was obtained although the sensitivity of the samples varied by a factor of 9. From these considerations it follows that the pH effect on x-ray sensitivity of the virus cannot be explained in terms of differing states of aggregation of the virus. A consideration of the charges on the suspended particles does provide an adequate explanation. The more acid the suspending medium, the greater will be the net positive charge on the suspended particles. The positively charged member of an ion pair produced by x-rays being of atomic size will be prevented from reaching the virus particle before losing its charge, while the negative ion being only of electronic mass will be able to penetrate to the virus and carry sufficient energy to inactivate the virus particle. The greater the net positive charge on the virus the greater will be its attraction for the free electrons produced by x-rays.

Conclusion.—1. Energy released by an ion pair produced within a virus particle may travel through that particle and inactivate one or more of the elementary virus units it may contain.

2. Electrons reaching the virus from the suspending medium have sufficient energy to inactivate the virus particle.

3. More of the electrons produced by x-rays in the suspending medium reach the virus particles when the latter carry a net positive charge than when the charge is negative. The greater the net positive charge the more electrons will be attracted.

* This work has been supported by the Columbia Foundation.

† Fellow of the Finney-Howell Cancer Research Foundation.

[1] Bawden, F. C., and Pirie, N. W., *Proc. Roy. Soc.* (*B*), **123,** 274 (1937).

[2] Frampton, V. L., *Science,* **95,** 232 (1942).

[3] Gowen, J. W., *Proc. Nat. Acad. Sci.*, **26,** 8 (1940).

[4] Holmes, Francis O., "Local Lesions in Tobacco Mosaic," *Bot. Gaz.*, **87,** 39 (1929).

[5] Lea, D. E., and Smith, K. M., *Parasitology,* **32,** 405 (1940).

[6] Marshak, A., *Jour. Gen. Physiol.*, **19,** 179 (1935); *Proc. Soc. Exp. Bro. & Med.*, **38,** 705 (1938); *Ibid.*, **39,** 194 (1938).

[7] Samuel, Geoffrey, and Bald, J. G., "On the Use of Primary Lesions in Quantitative Work with Two Plant Viruses," *Ann. Appl. Biol.*, **20,** 70 (1933).

[8] Smith, K. M., and MacClement, W. D., *Parasitology,* **32,** 320 (1940).

[9] Wyckoff, R. W. G., *Jour. Biol. Chem.*, **122,** 239 (1937).